

(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11)

EP 0 912 015 A2

(12)

## EUROPEAN PATENT APPLICATION

(43) Date of publication:  
28.04.1999 Bulletin 1999/17

(51) Int Cl. 6: H04L 12/26, H04L 12/56,  
H04L 12/28

(21) Application number: 98308309.8

(22) Date of filing: 13.10.1998

(84) Designated Contracting States:  
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE  
Designated Extension States:  
AL LT LV MK RO SI

(30) Priority: 14.10.1997 US 61790 P  
12.03.1998 US 77741 P  
22.05.1998 US 83759

(71) Applicant: LUCENT TECHNOLOGIES INC.  
Murray Hill, New Jersey 07974-0636 (US)

(72) Inventor: Chuah, Mool Choo  
Eatontown, New Jersey 07724 (US)

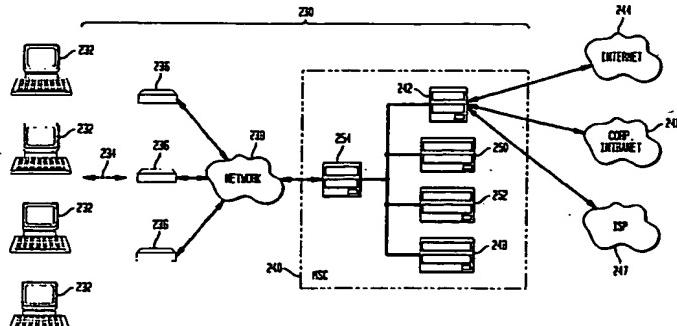
(74) Representative:  
Watts, Christopher Malcolm Kelway, Dr. et al  
Lucent Technologies (UK) Ltd,  
5 Mornington Road  
Woodford Green Essex, IG8 0TU (GB)

### (54) Method for overload control in a multiple access system for communications networks

(57) In the method for overload control in a wireless communications network employing On-Demand Multiple Access Fair Queuing, if the downlink/uplink buffer occupancy of the network has exceeded a high threshold, the base station determines if this is caused by a specific remote host or by a group of remote hosts. If caused by a specific remote host, the base station normally sends a flow control signal to the remote host to prevent it from sending more data, but may alternatively elect to disconnect other remotes if the remote experiencing bad performance is of a higher priority. The base station may additionally reduce the bandwidth shares allocated to any remote that have indicated tolerance for a variable allocated bandwidth. If the measured frame error rates for many remote hosts are increasing, then the base station may elect to disconnect those remote hosts that permit service interruption in order that more bandwidth may be allocated to the remaining us-

ers. If a majority of all associated remote hosts experience high uplink frame error rates, the base station may instead send a signal to a wireless hub which can coordinate the actions of other access points. Short packets queued up for so long at the base station that they exceed the time-to-live value allocated will be thrown away. The base station may also or alternatively elect to disconnect some users of a lower priority or redirect them to other nearby base stations that have a lower load. In a particular embodiment, an uplink Frame Error Rate (FER), an average uplink bit rate, a burstiness factor of uplink traffic, and a packet loss rate are measured at the base station for each remote host. Similarly, a downlink Frame Error Rate is measured at each remote host and then each FER is sent to the base station. If an overload condition exists, connections with a Frame Error Rate that has exceeded a threshold for a specified time and that have indicated that their connections can be interrupted are disconnected.

FIG. 2



applications. It is desired to extend such services to wireless networks. Research on merging ATM and wireless networks is therefore currently underway in many institutions and research laboratories. Many fundamental issues, affecting everything from the access layer to the transport layer, are being studied. Besides use of ATM as a transmission format at the air interface of a wireless network, ATM is also being considered for the wired infrastructure of cellular systems. Such a wired ATM infrastructure would be capable of supporting multiple access air interface technologies (e.g., CDMA, TDMA, etc.).

[0008] In a wireless network that supports multimedia traffic, an efficient channel access protocol needs to be maximize the utilization of the limited wireless spectrum while still supporting the quality of service requirements of all traffic. Several well-known channel access protocols are currently used in wireless data systems, such as Slotted Aloha, PRMA, etc.. Slotted Aloha is a simple protocol but, because it does not attempt to avoid or resolve collisions between data users, its theoretical capacity is just 0.37. In addition, Slotted Aloha is unsuitable for efficient transmission of variable-length packets.

[0009] Reservation-based protocols attempt to avoid and resolve collisions by dynamically reserving channel bandwidth for users needing to send packets. Typically, in such protocols a channel is divided into slots which are grouped into frames of N slots. A slot can be further subdivided into k minislots. Normally,  $N_1$  of the slots will be used for reservation purposes while the remaining  $N-N_1$  slots are data slots. The users that need to send packets send a reservation request packet in one of the  $M=N_1 \cdot k$  minislots. If the reservation request packet is successful, then the user will be allocated a certain number of data slots until the user or the base station releases the reservation. If the reservation request packet is not successful, the user will use a conflict resolution method to retransmit the reservation request until it is successfully transmitted.

[0010] A multiple access protocol for hybrid fiber-coax networks has been proposed by Doshi et al. in "A Broadband Multiple Access Protocol for STM, ATM, and Variable Length Data Services on Hybrid Fiber-Coax Networks," Bell Labs Technical Journal, Summer 1996, pp. 36-65. While sharing many issues with the wireless environment, this protocol does not completely address the unique problems encountered in the design of a wireless access scheme, such dealing with retransmissions over an error-prone wireless link and establishment of the transmission power level needed to ensure proper packet delivery. While this scheme does propose the idea of contention reservation slots, it does not provide a flexible scheme wherein the number of contention slots can be varied dynamically based on queue size information.

[0011] Karol et al have proposed a "Distributed-Queuing Request Update Multiple Access" scheme (DQRUMA) [Karol et al, "An efficient demand-assignment multiple access protocol for wireless packet (ATM) networks," Wireless Networks 1, pp. 267-279, 1995]. This wireless access scheme does not allow new users to contend for bandwidth during the conflict resolution period or utilize the reservation slot contention success rate during the previous round to adjust backoff time. This scheme also does not utilize a fair queuing technique, and hence does not make use of service tags to fairly allocate bandwidth between competing sources.

[0012] An important topic in designing a channel access protocol is selection of the scheduling techniques used to set the transmission order of uplink and downlink packets. A number of schedulers which are all variations on fair queuing have been proposed for wired networks [See, e.g., S.J., Golestani, "A Self-Clocked Fair Queuing Scheme For Broadband Applications", Proceedings of IEEE Infocom, 1994; Parekh and Gallagher, "A Generalized Processor Sharing Approach To Flow Control In Integrated Services Networks: The Single Node Case", IEEE/ACM Transactions On Networking, 1(3):344-357, June 1993; L. Chang, "Virtual Clock Algorithm", Proceedings of ACM Symposium, pp 1224-1231, 1992]. These all have the effect of providing access to a share of bandwidth as if each service class has its own server at its given rate.

[0013] The Weighted Fair Queuing scheme of Parekh and Gallagher is difficult to implement, so the Self-Clocked Fair Queuing (SCFQ) scheme was proposed by

$$F_k^i = \frac{L_k^i}{r_k} + \max (F_k^{i-1}, \bar{n}(a_k^i)) \quad (1)$$

[0014] Golestani . For SCFQ, the service tag is computed as;  
where  $\hat{u}(t)$  is the service tag of the packet in service at time t,  $F_k$  is the service tag for the  $i^{th}$  packet from class k with  $F_k^0 = 0$  for all k,  $L_k^i$  is the length of the  $i^{th}$  packet of class k,  $r_k$  is the relative weight assigned to class k, and  $a_k^i$  is the arrival time of the  $i^{th}$  packet of class k. Packets are then served in the order of these tag values. The algorithm of Golestani is designed for wired networks, however, and must be modified if it is to function in a wireless environment.

[0015] In particular the algorithm of Golestani does not address either how to handle transmission scheduling when the server (base station) does not have complete information about the size of the queues because they are remotely located or how to handle retransmission of lost packets.

[0016] Lu et al (University of Illinois) have proposed an "Idealized Weighted Fair Queuing" algorithm [Lu et al, "Fair

the following figures, wherein:

- Fig. 1 is a schematic diagram of a prior art network;
- 5 Fig. 2 is a schematic diagram of a network according to an aspect of the present invention;
- Figs. 3 and 4 are frame diagrams showing example downlink and uplink frame structures for a frequency division half-duplex embodiment of the invention;
- 10 Fig. 5 is a frame diagram of the synchronized downlink and uplink frame structures for a frequency division full-duplex embodiment of the invention;
- 15 Fig. 6A illustrates a frame having a general MAC layer downlink broadcast subframe, according to an example embodiment of an aspect of the present invention;
- Fig. 6B depicts a broadcast or multicast downlink frame format;
- 20 Fig. 6C depicts a beacon message format for the embodiment of Fig. 6B;
- Fig. 6D depicts a transmit permit format for the embodiment of Fig. 6B;
- 25 Fig. 6E depicts a transmit schedule format for the embodiment of Fig. 6B;
- Fig. 6F depicts a broadcast or multicast payload format for the embodiment of Fig. 6B;
- Fig. 7A illustrates a frame having a downlink unicast sub-frame, according to an example embodiment of the present invention;
- 30 Fig. 7B illustrates a flow control frame format for a downlink unicast data subframe, according to an example embodiment of the present invention;
- Fig. 7C illustrates a data frame format for a downlink unicast data subframe, according to an example embodiment of the present invention;
- 35 Fig. 7D depicts a unicast sub-frame concatenated to the back of a broadcast sub-frame;
- Figs. 8A illustrates a frame format for an uplink transmission frame, according to an example embodiment of the present invention;
- 40 Fig. 8B depicts the asynchronous transfer region of the frame of Fig. 8A;
- Fig. 8C depicts an uplink frame having reservation minislots according to an embodiment of the present invention;
- 45 Fig. 8D illustrates a frame format for a reservation minislot, according to an example embodiment of the present invention;
- Fig. 8E illustrates a frame format for a pure acknowledgment uplink frame, according to an example embodiment of the present invention;
- 50 Fig. 8F illustrates a frame format for a pure data uplink unicast frame, according to an example embodiment of the present invention;
- Fig. 8G illustrates a frame format for a combined acknowledgment and data uplink frame, according to an example embodiment of the present invention;
- 55 Fig. 8H illustrates a frame format for a combined acknowledgment, data, and & more uplink frame, according to an example embodiment of the present invention;

Detailed Description of Preferred Embodiments

[0023] As previously discussed, it is an object of the present invention to provide a wireless packet-switched data network for end users that avoids the public switched telephone network and provides end users of the wireless network with remote roaming capability. These and other objects are achieved in a wireless data network that includes a home mobility switching center, a foreign mobility switching center, a base station (access point) and an end user. The home mobility switching center includes a home registration server and a home inter-working function. The foreign mobility switching center includes a serving registration server and a serving inter-working function. The base station includes a proxy registration agent. The end user modem includes a user registration agent. The user registration agent is coupled to the proxy registration agent, the proxy registration agent is coupled to the serving registration server, and the serving registration server is coupled to the home registration server.

[0024] The proxy registration agent includes a module for sending an advertisement containing a care-of-address upon receipt of a solicitation from the user registration agent. The user registration agent includes a module for incorporating user identity information and the care-of-address into a registration request upon receipt of the advertisement, as well as a module for sending this registration request to the proxy registration agent. The proxy registration agent further includes a module for forwarding to the serving registration server any registration request received from any user.

[0025] The serving registration server includes a foreign directory module for determining a home registration server address, a module for encapsulating the registration request and incorporating serving registration server identity information and the encapsulated registration request into a radius access request when the home registration server address is determined, and a module for sending the radius access request to the home registration server. The home registration server includes a home directory module for authenticating the serving registration server identity information, a module for forming an inter-working function (IWF) request from the radius access request when the serving registration server identity information is authenticated, and a module for sending the inter-working request to the home inter-working function.

[0026] As seen in the embodiment of a network utilizing the present invention depicted in Fig. 2, end systems (remote hosts) 232 (for example, a portable Windows 95 personal computer) connect to wireless network 230 via external or internal modems. These modems allow end systems 232 to send and receive medium access control (MAC) frames over air link 234. If used, an external modem may be attached to PC or other end system 232 via a wired or wireless link. External modems are generally fixed, and could be co-located with roof top-mounted directional antennae. External modems may be connected to the user's PC using any appropriate linking method, including any of following types of links: universal serial bus, parallel port, infra-red, 802.3, or even an ISM radio link. Internal modems send and receive MAC frames over the air link and are preferably PCMCIA cards that are plugged into the laptop's backplane using a small omni-directional antenna.

[0027] Wide-area wireless coverage is provided by base stations (access points) 236. The range of coverage provided by base stations 236 depends on factors like link budget and capacity. Base stations are typically installed in cell sites by personal communication services (PCS) wireless service providers. Base stations 236 multiplex end system traffic from their coverage area to the system's mobile switching center (MSC) 240 over wire line or wireless microwave backhaul network 238.

[0028] At mobile switching center 240, packet data inter-working function (IWF) 252 terminates the wireless protocols for this network. IP router 242 connects MSC 240 to public internet 244, private intranets 246, or to internet service providers 247. Accounting and directory servers 248 in MSC 240 store accounting data and directory information. Element management server 250 manages the equipment, which includes the base stations, the IWFs, and the accounting/directory servers 248. The accounting server 248 collects accounting data on behalf of users and sends the data to the service provider's billing system. In a preferred embodiment, the interface supported by the accounting server 248 sends the accounting information in American Management Association (AMA) billing record format over a TCP/IP (transport control protocol/internet protocol) transport to a billing system (not shown in Fig. 2).

[0029] In the typical wireless network in which the present invention is utilized, each cell has a base station and a number of remote hosts (nodes), with or without additional wired hosts. Remote hosts/nodes can include any device capable of communication with the base station over a wireless link. Fixed-length packets arrive at the remote hosts ("remotes") at either a constant rate (CBR traffic) or according to various bursty random processes. The packets are buffered at the remotes until they are transmitted uplink to the base station, according to the channel access scheme. The base station broadcasts downlink packets that are destined for one or more of the remotes within its cell. Uplink and downlink communications are time-multiplexed on a single frequency channel in order to allow dynamic sharing of uplink and downlink bandwidths. The scheme of the invention can also be used for frequency division half-duplex (FDHD) and frequency division full duplex (FDFD) systems. The base station uses a variant of the Self-Clocked Fair Queuing algorithm of Golestani for scheduling the order of packet transmission from both remote hosts (remote queues) and wired hosts (local queues).

- [0038] As seen in Fig. 3, the downlink frame for the FDHD scheme of the invention may include physical layer overhead, such as some combination of guard and/or preamble bits 310 (which may be used as synchronizing bits), a medium access control (MAC) header 312, various control messages such as certain types of beacon messages 314, transmit permits 320, minislot information for the next uplink frame 350, and transmit schedules 322, acknowledgments (ACKs) for the reservation of minislots in previous uplink frame 330, acknowledgments for the data sent in the previous uplink frame 340, broadcast/multicast data messages 360, unicast data messages 380, and a frame check sequence (FCS) 355 for each preceding data message. Not all fields and messages are necessarily found in each downlink frame. For example, a downlink frame may consist of just the transmit permits, acknowledgments for reservation minislots, and unicast messages.
- [0039] Some control messages are preferably part of the broadcast message 360, which may include such things as load metric, information about reservation minislots, flow control information, acknowledgments, and power management parameters. The load metric information can be as simple as the number of remote nodes registered with the AP, or may be more sophisticated, such as the equivalent number of active remote nodes. The load metric can be used for admission control and load-balancing among APs. The minislots information describes the number of reservation minislots present in the next uplink frame, if any, and their locations. The flow control information contains the connection cookie (identity) and an Xon/Xoff indication.
- [0040] The acknowledgment 340 for uplink unicast traffic can be as simple as acknowledgment bits that are part of the broadcast message, or may be more sophisticated, such as separate unicast messages which specify the connection identity and the sequence number of the message to be acknowledged. In the former case, if the uplink transmission uses a frame structure with N fixed basic slots, then at most only N acknowledgment bits are needed. For the latter case, it is necessary for each message to have a separate frame check sequence (FCS). Note that, due to the "hidden terminal problem," all the frames transmitted need to be acknowledged.
- [0041] The data slots 380 include transmissions from multiple remote nodes. The transmission from each remote node includes guard bits, preamble bits, frame control bits, acknowledgments, and/or data messages. One of the frame control bits is a "more" bit that is used to indicate that the remote node has more data to transmit. Alternatively, the number of remaining bytes or number of fixed size packets left to be transmitted may be particularly specified, rather than just through use of a "more" bit.
- [0042] As seen in Fig. 4, the FDHD uplink frame generally will consist of a contention period 410 and a contention-free period 415. The contention period 410 includes one or more contention slots, each of which can be either a contention data slot 420 or a contention reservation slot 422. The contention-free period 415 consists of acknowledgments 440 for previous downlink dataslots and multiple data slots 480 and 486. If desirable, these contention slots 420 and 422 may be spread uniformly across the whole frame rather than clustered together. Each contention reservation slot 422 may be further subdivided into k subslots 430, called reservation minislots. Each minislot 430 is long enough to contain the identity of a remote node, generally around 30 bytes. Contention slots 420 may be utilized as dataslots for transmitting small data packets. The contention-free period 415 may include pure ACK frames 440, pure data frames 480, and/or combination frames 486 having both data 488 and ACK 490 portions.
- [0043] The number of minislots 430 may be dynamically changed. If, for example, there are k minislots in a contention reservation slot 422 and N total contention slots, N1 of which are reservation slots 422 containing a total of N1\*k minislots, then the remaining (N-N1) slots are currently contention data slots. If there are a minimum and maximum number of reservation minislots desired for the system, the number of available reservation minislots can be dynamically changed based on the percentage of idle minislots and the total uplink queue length. Several methods for dynamically changing the number of minislots are described later in conjunction with Figs. 12A-12D.
- [0044] In order to assign different priorities to the remote nodes attempting to gain access to the system, the  $M_1 = N_1 * k$  minislots (where  $N_1$  is the number of contention reservation slots) may be divided into various groups. For example, a group of remote nodes with MAC addresses within a certain range may only be allowed to randomly access up to  $M_2$  minislots (where  $M_2 < M_1$ ), whereas a higher priority group of remote nodes with MAC addresses within another range may be allowed to randomly access up to  $M_1$  minislots. Alternatively, priority classes may be assigned to nodes based on connection identity rather than MAC address. A priority assignment feature could be particularly useful, for example, for emergency-response organizations, such as hospital or police staff, and could be achieved through the provision of wireless modems that have a higher priority of access than regular wireless modems. This feature could also be sold as a service class to customers who are willing to pay more for a higher access priority.
- [0045] As depicted in Fig. 5, uplink frames 502 and 512 in Frequency Division Full-Duplex (FDFD) mode are synchronized with the downlink frames 562 and 572. As seen in Fig. 5, uplink frames 502 are shown as viewed from the wireless modem, uplink frames 512 are shown as viewed from the AP, downlink frames 562 are shown as viewed from the AP, and downlink frames 572 are shown as viewed from the wireless modem. In Fig. 5, the AP has previously sent downlink frame n to the wireless modem, which has received it after a propagation delay  $T_p$ . In response, after end system processing time  $T_{cpe}$ , the wireless modem sends uplink frame n 504, which is received 514 by the AP at propagation delay  $T_p$  520 later. Meanwhile, the AP has already begun transmission of downlink frame n+1 564.

the remote node or connection is allowed to transmit 659 (end slot). In the example depicted, the Message Length 655 is 6 bytes, meaning there are two transmit permits 656 following. The first transmit permit 656 is for remote node 657 #3, which may start transmission at start slot 658 #1 and may transmit through end slot 659 #2. The second transmit permit 656 is for remote node 657 #5, which may start transmission at start slot 658 #3 and may transmit through end slot 659 #5. Different "Type" and "Subtype" labels may be used for the transmit permits of those wireless modems to which the AP sends both downlink unicast data and transmit permits. Subframes combining transmit permits and schedules are preferably sent after the pure transmit permits and before any pure transmit schedules.

[0054] Fig. 6E depicts the transmit schedule format of the embodiment of Fig. 6B. The optional transmit schedules 661 (322, Fig. 3) allow remote nodes or connections that are associated with the AP to power down if no more data is scheduled to be sent to them. The transmit schedule body 661 is preceded by type 662 "Control" and subtype 664 "Transmit Schedule" fields. The transmit schedules 661 can take one of two forms. The first form is simple, e.g. a bitmap having a "1" to indicate the presence of unicast data for that remote node or connection, so that, for example "011000000010" would indicate that the frame contains unicast data for the second, third, and eleventh of twelve remote nodes. The second possible form is more sophisticated, containing, for example, a remote node or connection ID, the start time, and the duration that the node is allowed to transmit (the same as the data contained in a transmit permit).

[0055] Fig. 6F depicts the broadcast or multicast payload format 670 (360, Fig. 3) of the embodiment of Fig. 6B. The payload body 671 can contain a wide variety of data messages or control information and is preceded by a type field 672 and a subtype field 674. These fields will vary according to the content of the payload body 671, for example if the payload body 671 contains the number of contention minislots and their positions, the type 672 is "Control" and the subtype 674 is "Contention Minislot Information," whereas if the payload body 671 contains a broadcast message from a wireless hub, type 672 will be "Data" and subtype 674 will also be "Data."

[0056] Fig. 7A depicts an example embodiment of a frame format of a downlink unicast sub-frame 700 according to the present invention. Examples of unicast subframes are control messages, such as association response frames and flow control request frames, and data messages, with acknowledgments and/or "more data" information. The "more data" information can be as simple as one bit in the Frame Control 702 subfield of the MAC header, or may be more particularly expressed as the number of remaining bytes to be transmitted. The example downlink unicast subframe 700 depicted in Fig. 7A has a MAC header 701 having a one-byte Frame Control subfield 702, a 2-byte Frame Duration field 704, a 6-byte Source MAC Address 706, a 6-byte Destination MAC Address 708, and a 2-byte Sequence Control field 710. The remainder of the downlink unicast subframe 700 is comprised of the unicast data body 720 and a frame check sequence (FCS) 712.

[0057] Fig. 7B depicts an example embodiment of a flow control frame format for a downlink unicast data sub-frame according to the present invention. In the particular embodiment of Fig. 7B, the unicast data body 720 has a Type field 722 "Control" and Subtype field 724 "Flow Control", followed by a Connection Identity (CC) field 726. Data field 730 follows, containing an Xon/Xoff bit.

[0058] Fig. 7C depicts an example embodiment of a data frame format for a downlink unicast data subframe according to the present invention. In the embodiment of Fig. 7C, unicast data body 720 contains one or more of the following fields: Data 744, ACK 746 and "More Data" 748. If present, More Data field 748 can be as simple as a 1-bit flag or may give the remaining number of bytes. ACK field 746, if present, may take the form of a sequence number or a bitmap. Data body 720 starts with a Type field 740 "Data" and a Subtype field 742 that can have the values "Data", "Data + ACK", "Data + ACK + More", or "ACK", depending on the composition of the fields following.

[0059] If there is only one connection per wireless modem, then unicast sub-frames may be concatenated so that they are attached to the back of a broadcast subframe without the cost of source MAC address field overhead, as shown in Fig. 7D. The frame of Fig. 7D is comprised of a unicast subframe 700 concatenated with a broadcast subframe 750. Broadcast subframe 750 is comprised of a 6-byte Source MAC Address 752, a 6-byte Destination MAC Address 754, a one-byte Frame Control subfield 756, a 2-byte Frame Duration field 758, a 2-byte Sequence Control field 760, a broadcast data field 762, and a frame check sequence (FCS) 764. Unicast subframe 700 is comprised of a 6-byte Destination MAC Address 708, a one-byte Frame Control subfield 702, a 2-byte Frame Duration field 704, a 2-byte Sequence Control field 710, Type field 722, Subtype field 724, Connection Identity 726, data field 730, and a frame check sequence (FCS) 712. Frame Control field 702 in the unicast subframe 700 is optional, generally being included if the bits in the Frame Control field can be expected to change frequently. If the Frame Control field of the unicast subframe can be expected to be relatively static, it will frequently be omitted except on the specific occasions it is required.

[0060] For synchronization purposes, the AP may schedule the downlink broadcast and unicast subframes in such a way that the total broadcast and unicast subframe transmission time falls within an x ms frame structure, where x is generally 2 ms. However, for uplink transmission, uplink communication from the wireless modem is in burst mode and subject to collision in any case where more than one modem transmits in a given time window. Such a collision can be detected only at the AP. Each transmission burst also necessarily involves some physical layer overhead.

[0061] To accommodate these factors, as shown in Fig. 8A, a frame structure has been defined for uplink transmission

off time 4 us. These parameters lead to the requirements of a guard time of 20 bits at each end of a physical layer PDU and a preamble of 64 bits. In this system, a 2 ms uplink frame corresponds to 640 bytes. Assuming that the frame consists of both an STR and an ATR and that each basic slot in the STR is 27 bytes long, then a frame with one STR slot can also have, for example, 10 reservation minislots (with each basic slot being converted to 5 reservation minislots), 2 data contention slots, and 5 reserved data slots for ATM PDUs or VL PDUs.

[0071] As illustrated in Fig. 11, a downlink broadcast/multicast message may be used for paging request messages. The paging request and associated response messages are designed to enable a PC on a wired network to call another PC over the wireless network. Paging request messages are useful for alerting a wireless modem that a wired host or another wireless modem is interested in communicating with it. The wireless modem whose ID is contained in a received paging request message responds with a paging response message, as well as with a connection request if there is currently no connection between the wireless modem and the Access Point. Paging capability requires a location server, which may be co-located with a PPP server if desired. The method would normally be used when the PC accessed via the wireless network has no IP address through which it may be more efficiently accessed.

[0072] As illustrated in Fig. 11, in order to allow PC2 1102 to initiate a call to PC1 1104 which is attached to a wireless modem 1106, a paging request message is defined. The initiating PC (PC2) 1102 sends a Call\_Initiate message 1110 to a location/PPP server 1112 which identifies the home registration server 1116. The home registration server 1116 then identifies the proper WH/IWF and relays 1118 the Call\_Initiate message to the AP 1120. Next, the AP 1120 sends a paging request 1130 to the wireless modem 1106 with which PC1 1104 is associated. Finally, the wireless modem 1106 relays 1132 the Call\_Initiate message to PC1 1104.

[0073] To accept the call, PC1 1104 sends a Call\_Accept message 1140 to the wireless modem 1106, simultaneously with a Connect\_Request message. The wireless modem 1106 then sends a paging response 1142 to the AP 1120, which relays 1144 the message to the WH/IWF 1116. The wireless modem 1106 also relays the Connect\_Request message to the AP 1120, which similarly relays it to the WH/IWF 1116. The WH/IWF 1116 sends a Connect\_Reply message 1145 to PC1 1104 and then relays a Call\_Accept message 1146 back to the location server 1112. Finally, the location server 1112 relays 1148 the Call\_Accept message to PC2 1102.

[0074] The ODMAFQ scheme is capable of providing priority access within the same message stream from each user. Priority access will generally give important control messages a higher priority than data messages. Some important control messages which might be transmitted by a wireless modem in a reservation slot include: (a) Association Request, for requesting association of the wireless modem with an Access Point, (b) Connect Request, for requesting a connection set-up, (c) Paging Response, for responding to a Paging Request, and (d) Bandwidth Request, for requesting bandwidth allocation after having been silent for a while. The various types of possible messages may also be assigned correspondingly different priorities for differing Qualities of Service. In general, Association Request, Connect Request, and Paging Response messages would be expected to have a higher priority than data messages. As an example, if the service provider wishes to admit more users, Bandwidth Request messages should then be given lower priority than Connect Request and Paging Response messages, allowing for faster connection set-ups. Among data messages, voice signals carried over RTP/UDP packets, for example, would generally be given higher priority than tcp/ip data packets.

[0075] A fragmentation/reassembly mechanism has been defined in order to allow for fragment retransmission. The AP and wireless modem will generally fragment the MAC layer service data unit (SDU) if it exceeds the maximum payload size or if it exceeds the remaining space available in a downlink or uplink frame. Alternatively, a fragmentation threshold may be defined beyond which the MAC SDU will be fragmented. Each fragment has a sequence control field. All fragments belonging to the same SDU carry the same 12-bit sequence number, but are assigned different fragment numbers. A "More Fragment" bit in the frame control field is then set for all fragments except the last, indicating that there are additional fragments still to follow. The fragments are then sent in order of lowest to highest fragment number.

[0076] To meet the in-sequence delivery requirement, both the AP and the wireless modem make sure that all the fragments of the same SDU are transmitted before a new SDU is transmitted. Only those fragments that are lost are retransmitted. To prevent endless transmission delay (with concomitant transmission backlog), a particular source (wireless modem or AP) maintains a MAC SDU transmission timer which is started the moment a MAC SDU is passed to the MAC layer. When the timer exceeds the pre-established MAC SDU lifetime, all remaining fragments will be discarded by the source, and no attempt is made to complete the transmission of the MAC SDU.

[0077] To prevent endless waiting for permanently lost fragments, the destination station reconstructs the MAC SDU by combining the fragments in order of the fragment number of the sequence control field. If the destination station receives a fragment with the "more fragment" bit set, it knows that it has not yet received a complete MAC SDU. As soon as the destination station receives a fragment having a clear "more fragment" bit, it will reassemble the MAC SDU and pass it to a higher layer.

[0078] The destination station (such as a wireless modem or AP) maintains a receive MAC SDU timer which is initiated upon receiving the first fragment of a MAC SDU. The destination station should preferably have at least 3

1. A random number,  $x$ , is generated at the remote node modem from a uniform distribution over 1 through M, and
2. The initial contention message is transmitted in the  $x$ th minislot in the next uplink frame.

5 [0085] If desired, carrier sensing can also be used during initial contention. Before transmission, the channel is sensed. If access priority is implemented, instead of choosing a random number between 1 and M, the wireless modem then chooses between 1 and  $I_i$  where  $I_i$  is the threshold for users of class i, where a lower value indicates a higher priority, i.e.,  $I_{i+1} < I_i$ . If, however, the contention message is not a contention reservation minislot request message, but rather is a contention data slot message, then the message is transmitted in the next contention data slot.

10 [0086] More than two access priority classes may be offered. As previously discussed, the uplink frame includes  $N_i$  minislots. If, for example, if there are  $p$  access priority classes, each class having access priority  $i$  (where a smaller number means a higher priority) can send contentions in the minislots ranging from 1 to  $I_i$ , where  $I_1 = N_1$ ,  $I_{i+1} \leq I_i$ . A strict usage priority can be implemented on top of this access priority scheme, so that when an AP receives a connection request that has a higher usage priority, it can disconnect an existing connection of a lower usage priority by sending a disconnect request frame to the wireless modem that supports the connection.

15 [0087] Collision occurs in a contention slot when two or more wireless modems transmit in the same minislot. Also, if interference causes corruption of data in a contention slot, the slot status is declared to be a COLLISION. As previously described, there are 2 types of contention slots in an uplink frame: (1) a reservation slot containing minislots for bandwidth request messages, and (2) a data slot containing uplink short bursty messages in contention superslots. At the 20 AP, the RF energy in an uplink contention time slot is estimated. If there is no energy present, then the contention slot is declared IDLE. The status of a contention slot is declared to be SUCCESS if all the following hold true: 1) RF energy has been detected in the slot, 2) a preamble in that slot is not corrupted, and 3) a frame check sequence (FCS) in the slot indicates no errors. The status of a contention slot is declared to be COLLISION if RF energy has been detected in the slot, and at least one of the following holds true: 1) the preamble in that slot is corrupted or 2) a frame check 25 sequence (FCS) in the slot indicates error.

[0088] Fig 18A illustrates is an embodiment of a method for access control according to an aspect of the present invention.  $N$  contention reservation minislots are configured in each uplink frame 1810. The  $N$  minislots are organized into a plurality of access priority classes, each class having a different priority. The AP is configured to allow  $m$  access priority classes 1815. Each remote host of access priority class  $i$ , randomly picks 1820 one contention minislot and transmits an access request, the contention minislot picked being in a range from 1 to  $N_i$  where  $N_{(i+1)} < N_i$  and  $N_1=N$ . The base station receives 1825 the access requests and sequentially examines the received contention minislots. If the minislot currently being examined contains an uncollided request 1830, the AP grants access 1835 to the remote host corresponding to the uncollided access request. If the minislot currently being examined contains a collided request 1830, the AP will not send an ACK, causing the affected remote nodes to perform conflict resolution 1840. After the 30 conflict resolution period, the AP grants access to the "winning" remote host 1845. Meanwhile, if more minislots remain to be examined 1850, the AP continues to check minislots for collisions 1830, either granting access to successful requesting hosts 1835 or awaiting the outcome of conflict resolution 1840.

[0089] Fig. 18B is a flowchart illustrating an alternate embodiment of a method for access control according to an aspect of the present invention. organized into a plurality of access priority classes, each with a different priority.  $N$  contention reservation minislots are configured in each uplink frame 1810. The  $N$  minislots are organized into a plurality of access priority classes, each class having a different priority. The AP is configured to allow  $m$  access priority classes 1815. Each remote host of access priority class  $i$  and with a stack level that equals 0, then transmits an access request with a probability  $P_i$  where  $P_{(i+1)} < P_i$  and  $P_1 = 1$  1860. The base station receives 1825 the access requests and sequentially examines the received contention minislots. If the minislot currently being examined contains an uncollided request 1830, the AP grants access 1835 to the remote host corresponding to the uncollided access request. If the minislot currently being examined contains a collided request 1830, the AP will not send an ACK, causing the affected remote nodes to perform conflict resolution 1840. After the 35 conflict resolution period, the AP grants access to the "winning" remote host 1845. If more minislots remain to be examined 1850, the AP continues to check minislots for collisions 1830, either granting access to successful requesting hosts 1835 or awaiting the outcome of conflict resolution 1840.

[0090] IDLE, SUCCESS and COLLISION status information is conveyed back to the wireless modems. The AP places the slot status information in the downlink reservation acknowledgment field. There are three alternative preferred conflict resolution methods that may be used. The first method is suggested in the IEEE 802.14 standard, and is described along with two new methods below. Simulation results show that the second method described provides 40 a better access delay.

[0091] In the first conflict resolution method, suggested in IEEE standard 802.14, each wireless node that wishes to transmit randomly picks one of the reservation minislots. If a collision is indicated, a modem that was affected by the collision retransmits based on a random binary-exponential back-off method. This backoff method operates in accord-

D. A wireless node that is in the request state and that did not send an access request (i.e., a node backlogged with stack level > 0) will decrement its stack level by one if the outcome reported in the reservation acknowledgment field for the assigned minislot is SUCCESS.

5 [0095] The operation of this method is depicted in Fig. 14B. A wireless node waiting to access the AP or send new data 1432 sets its stack level to 0 and enters the request state. If the stack level of the node is 0 1434, the node randomly picks 1436 a reservation minislot for transmission of an access request and transmits the access request. If the outcome of the request is SUCCESS 1438, and the queue at the node is empty 1439, the node transmits 1440 the current packet and exits the request state, returning to the waiting state 1432. If the queue at the node is not empty 1439, then, after receiving a transmit permit from the AP, the node transmits 1441 the current packet along with a piggybacked reservation request for transmission of the next packet in its queue, continuing to transmit packets with piggybacked reservation requests 1441 after receiving transmit permits until the queue is empty 1439, at which point it transmits the remaining packet 1440, exits the request state, and returns to the waiting state 1402.

10 [0096] If the outcome of the reservation request 1436 was not SUCCESS 1438, the node participates in a random draw 1444 to learn whether to increment 1448 its stack level by 1 or leave 1446 its stack level at 0. If the stack level remains 1446 at 0, the node again randomly picks 1436 a reservation minislot for transmission of an access request and transmits the access request. If the stack level is incremented 1448, the stack level will not be 0 1434. If the stack level of any remote node is not 0 1434 then if the outcome of the previous reservation request was COLLIDED 1450, the node increments 1452 its stack level by 1. If the outcome for the previous reservation request was not COLLIDED 1450, the node decrements 1454 its stack level by 1.

15 [0097] The third conflict resolution method is a modification of the second. In the third conflict resolution method, the modem in each wireless node is again characterized by a stack level, and only wireless nodes with a stack level equal to zero are permitted to transmit access request packets. Modems with stack level greater than zero are regarded as backlogged. The rules of the third method are:

- 25 1. When a wireless node first wishes to gain access to the network or has gained access and wishes to send new data, it is placed in a request state and assigned a stack level of zero.
- 30 2. When there are M reservation minislots, each wireless node in a request state randomly picks one of the M reservation minislots to be its assigned minislot in which to transmit an access request packet.
- 35 3. When the wireless node is characterized by a stack level equal to zero, it transmits an access request packet; however, when the remote node is characterized by a stack level other than zero, it does not transmit an access request packet.
- 40 4. At the end of the time slot, each wireless node changes its stack level based on the outcome (either COLLIDED, IDLE or SUCCESS) of all access requests as reported in the reservation acknowledgment fields of a downlink message from the Access Point.
- 45 A. A wireless node that sent an access request and received a SUCCESS outcome will be removed from the request state.
- B. A wireless node that sent an access request and received a COLLIDED outcome will either increment its stack level by one or leave its stack level at zero depending on the outcome of a random draw.
- C. A wireless node that is in the request state and that did not send an access request (i.e., a node backlogged with stack level > 0) will decrement its stack level by one if the outcomes of all access requests reported in at least 80% (or some other predefined threshold) of the reservation acknowledgment fields is either SUCCESS or IDLE. Otherwise, the remote node will increment its stack level by one.
- 50 D. When the backlogged modem's stack level is decremented to zero, the modem randomly picks one of the M minislots (or the  $l_i$  minislots if access priority is implemented) to resend its request.

55 [0098] The operation of this method is depicted in Fig. 14C and is similar to that of the method of Fig. 14B. A wireless node waiting to access the AP or send new data 1432 sets its stack level to 0 and enters the request state. If the stack level of the node is 0 1434, the node randomly picks 1436 a reservation minislot for transmission of an access request and transmits the access request. If the outcome of the request is SUCCESS 1438, and the queue at the node is empty 1439, the node transmits 1440 the current packet and exits the request state, returning to the waiting state 1432. If the

```

If((q > HIGH_THRESH) && (idle > IDLE_THRESH1)) {

    5      If (State!=1) {

        no_mini=no_mini-k;
        no_slots=no_slots+1;

    10     State=1

        }

    15     }

If ((q < LOW_THRESHOLD) && (idle < IDLE_THRESH2)) {

    20     If (State==1) {

        no_mini=no_mini+k;
        no_slots=no_slots-1;

    25     State=0

    30     }

}

```

35 [0104] As shown in Fig. 12A, if the total uplink queue length is greater than a high threshold (HIGH) 1201, then if the  
 percentage of idle minislots (IDLE) is not greater than a first idle threshold (IDLE1) 1202, the number of minislots (N)  
 is left unchanged. If, however, the percentage of idle minislots is greater than the first idle threshold 1202, and the  
 state is not "1" 1203 (meaning that the number of minislots was not just decreased), the number of minislots in the  
 frame is decreased 1204 by some k, the number of dataslots (SLOTS) in the frame is increased by 1, and the state is  
 set to "1". If the total uplink queue length is not greater than the high threshold 1201, then if the total uplink queue  
 length is less than a low threshold (LOW) 1205, and the percentage of idle minislots is not less than a second idle  
 threshold (IDLE2) 1206, the number of minislots is left unchanged. If, however, the percentage of idle minislots is less  
 than the second idle threshold 1206, and the state is "1" 1207 (meaning that the number of minislots was just de-  
 creased), the number of minislots in the frame is increased 1204 by k, the number of dataslots in the frame is decreased  
 by 1, and the state is set to "0". In all four methods, the threshold values and the value of k may be prespecified if desired.  
 40 [0105] A software implementation of method 2 for the dynamic adjustment of the number of reservation minislots is  
 given below and is also illustrated pictorially in the flowchart of Fig. 12B. In the methods of Fig. 12B and 12D, HIGH2  
 >HIGH1 and LOW2>LOW1.

50

55

```
else if ((q > HIGH1) && (idle > IDLE_THRESH1)) {  
    If (State==0) {  
        5      no_mini=no_mini-k;  
        no_slots=no_slots+1;  
        State=1  
    }  
    }  
    }  
    15  
If ((q< LOW1) && (idle < IDLE_THRESH2)) {  
    If (State> 0) {  
        20     If (State==1) {  
            no_mini=no_mini+k;  
            no_slots=no_slots+1;  
            State=0;  
        }  
        }  
        25     else {  
            no_mini=no_mini+2k;  
            no_slots=no_slots-2;  
            State=0;  
        }  
        }  
        30  
    }  
    }  
    35  
else if ((q < LOW2) && (idle < IDLE_THRESH2)){  
    If (State==2) {  
        40      no_mini=no_mini+k;  
        no_slots=no_slots-1;  
        State=1  
    }  
    }  
    45
```

```

5      If((q > HIGH_THRESH) && (idle > IDLE_THRESH1)) {

10     If (no_mini > NUM_MINI_MIN) {
15       no_mini=no_mini-k;
20       no_slots=no_slots+1;
25       }

30     If ((q < LOW_THRESHOLD) && (idle < IDLE_THRESH2)){
35       If (no_mini< NUM_MINI_MAX) {
40         no_mini=no_mini+k;
45         no_slots=no_slots-1;
50         }

55   }
}

```

[0111] As shown in Fig. 12C, if the total uplink queue length is greater than a high threshold 1240, then if the percentage of idle minislots is not greater than a first idle threshold 1241, the number of minislots is left unchanged. If, however, the percentage of idle minislots is greater than the first idle threshold 1241, then if the number of minislots is greater than the minimum number of minislots allowed (MIN) 1242, the number of minislots in the frame is decreased 1243 by k and the number of dataslots in the frame is increased by 1. If the total uplink queue length is not greater than a high threshold 1240, then if the the total uplink queue length is less than a low threshold 1244, and the percentage of idle minislots is not less than a second idle threshold 1245, the number of minislots is left unchanged. If, however, the percentage of idle minislots is less than the second idle threshold 1245, and the number of minislots is less than the maximum number of minislots allowed (MAX) 1246, the number of minislots in the frame is increased 1247 by k and the number of dataslots in the frame is decreased by 1.

[0112] A software implementation of method 4 for the dynamic adjustment of the number of reservation minislots is given below and is also illustrated pictorially in the flowchart of Fig. 12D.

[0113] As shown in Fig. 12D, if the total uplink queue length is greater than a first high threshold 1250, then if the percentage of idle minislots is not greater than a first idle threshold 1251, the number of minislots is left unchanged. If, however, the percentage of idle minislots is greater than the first idle threshold 1251, and the number of minislots is greater than a minimum number of minislots allowed 1252, the number of minislots in the frame is decreased 1253 by 2k and the number of dataslots in the frame is increased by 2. If the total uplink queue length is not greater than the first high threshold 1250, then if the total uplink queue length is greater than a second high threshold 1254, and the percentage of idle minislots is not greater than the first idle threshold 1255, the number of minislots is left unchanged. If, however, the percentage of idle minislots is greater than the first idle threshold 1255, and the number of minislots is greater than the minimum number of minislots allowed 1256, the number of minislots in the frame is decreased 1257 by k and the number of dataslots in the frame is increased by 1.

[0114] In the method of Fig. 12D, if the total uplink queue length is not greater than the first high threshold 1250 and the second high threshold 1254, but is also not lower than both a first 1258 and second 1262 low threshold, the number of minislots is left unchanged. If, however, the total uplink queue length is not greater than the second high threshold 1254, is not lower than the first low threshold 1258, but is lower than the second 1262 low threshold, then if the percentage of idle minislots is less than a second idle threshold 1263, and the number of minislots is less than the maximum number allowed 1264, the number of minislots in the frame is increased 1265 by k and the number of dataslots in the frame is decreased by 1.

[0115] If the total uplink queue length is not greater than the second high threshold 1254 and is lower than the first low threshold 1258, then if the percentage of idle minislots is less than the second idle threshold 1259, and the number of minislots is less than the maximum number allowed 1260, the number of minislots in the frame is increased 1261 by 2k and the number of dataslots in the frame is decreased by 2.

[0116] The role of the AP in responding to uplink bandwidth requests from modems, whether they arrive in pure reservation minislots or in piggybacked form, is to control uplink transmission in order to achieve a balance between high bandwidth efficiency and excellent quality of service (QoS) management. While QoS requirements for constant bit rate CBR traffic are extremely important and stringent, they are relatively liberal for traditional data traffic. One goal of the bandwidth allocation scheme in the AP is therefore to take advantage of these diverse QoS requirements in order to achieve a high degree of statistical multiplexing. In order to determine how the AP should transmit downlink traffic from various connections, the AP requires a downlink scheduling system. Similarly, in order to coordinate the uplink transmissions from associated wireless modems, the AP requires a system for scheduling the uplink transmission opportunity of each wireless modem. The scheduling systems can be as simple as round-robin, strict priority, or a first come-first serve algorithm, or may alternatively be more complex, such as a fair queuing algorithm. As discussed previously, a number of schedulers which are all variations on fair queuing have been proposed.

[0117] The uplink scheduling system is not required to be the same as the downlink scheduling system, however, for a simple embodiment, one may elect that they be the same. Obviously, a scheduling system is desired that provides Quality of Service to end users. As in ATM networks, different service classes can be defined to cater to the diverse QoS needs of different applications. The possible service classes include: constant bit rate (CBR), real-time and non-real-time variable bit rate (RT VBR, NRT VBR), unspecified bit rate (UBR), and available bit rate (ABR). In order to meet the QoS requirements of different service classes, there needs to be a method for allocation of bandwidth and buffer resources that does not require statically prioritizing one over the other.

[0118] In order for the AP to perform downlink and uplink scheduling in the case where the wireless modems are geographically distributed, a mechanism is needed for the wireless modems to pass relevant information to the base station, which is the only location that has a complete view of all transmission queues (i.e., transmission queues for both wired and wireless hosts). There are at least two alternative ways to compute the service tags for all hosts associated with the access point. In these methods, the wired hosts, with which the associated wireless modems are communicating, are assumed to be permanently associated with the access point. In one method, the base station can broadcast the system virtual time and the assigned shares of service classes to each of the wireless modems. Then, each wireless modem computes its own service tag and informs the base station of it via a request access packet or by piggybacking on the data transmission. Alternatively, the wireless modem can simply inform the base station of its queue size (again via a request access packet or by piggybacking on data transmission), and the base station can compute the service tag for each wireless modem as well as for the wired hosts. The second method is more efficient in terms of downlink bandwidth utilization, since the base station does not have to transmit the assigned service shares (which may be dynamically varying) to each wireless modem.

[0119] An embodiment of the first method is illustrated in Fig. 15A. The base station broadcasts a system virtual time 1510 to the remote hosts. Each remote host computes a service tag value 1515 for each of its newly arrived packets, then transmits 1520 the smallest tag value to the base station. Transmit permits are then assigned 1530 at the base station based on the service tag values received from the remote hosts and the available data slots. The transmit permits are broadcast to the remote hosts 1540, and then packets are received from the remotes 1545 in the order specified by the transmit permits. If a packet is lost or is received having errors 1545, the sending remote is made

and 32.

[0125] FIG. 9A shows the service tags of the packets of this example at time t = 0. The packets of session 1 with service tag 8 902, service tag 16 904, service tag 24 906, and service tag 32 908 are interleaved with packets 912, 914, 916, 918, 920, 922, and 924 from session 2. Packet 910 from session 2, having service tag 4, is currently in service.

[0126] FIG. 9B shows the service tags of the remaining queued packets at time t = 3, just before the packets from session 3 arrive. Packet 918 from session 2, having service tag 20, is currently in service. FIG. 9C shows the service tags of the packets at time t = 3, just after the 9 packets 930, 932, 934, 936, 938, 940, 942, 944 and 946 from session 3 arrive. Note that the service tag for the first packet 930 of session 3 starts at 22 because, when the packets arrive, the service tag of the packet currently being served was 20. Thus, for a service tag increment of 2, the first packet 930 from session 3 will receive service tag 22. Subsequent packets from session 3 then have service tags of 24, 26, 28, etc..

[0127] FIG. 9D shows the service tags of the remaining queued packets at time t = 4.5. The transmission of the packet 906 with service tag 24 from session 1 has errors. The access point therefore recomputes a new service tag of 32 for this packet 950, which needs to be retransmitted. The access point also recomputes the service tags of the remaining packets from session 1, which in this case only affects one other packet 952 (908 in Fig. 9C), which receives a new service tag of 40. In this way, the retransmission of a packet from a particular session does not affect the Quality of Service of other sessions.

[0128] When the remote host PC wants data services, it sends a connect message to the wireless modem. Upon receiving this message, the wireless modem monitors the broadcast frame that is being continuously sent by the AP. The beacon message is part of this broadcast frame and provides timing information, the ESS-ID of the Network, the BSS-ID of the AP, information about the contention slots, the load metric of the AP, etc.. The wireless modem then chooses the AP with which it wants to associate and sends a MAC layer associate request frame. Since association request frames are sent in contention mode, collisions may occur. A wireless modem needs to retransmit the association request frame if it does not receive an association response frame from the AP. After a maximum number of retries, the wireless modem will send a connect fail message to the remote host PC, indicating that the wireless modem cannot associate with an AP at this time.

[0129] Upon receiving an associate request frame from a wireless modem, after the AP has successfully authenticated the wireless modem, it sends an association response frame with a status code "successful" to the modem. Authentication is performed at the network layer. When a user requests a connection via the wireless modem, the connection request is forwarded by the access point to the wireless hub. The wireless hub then authenticates the user. If the user is successfully authenticated, a unique connection cookie is provided by the wireless hub to the access point. If it is desirable to provide different QoSs to different connections from the same user, then different connection cookies are assigned to the same user; similarly, if it is desirable to provide different QoSs to different users (albeit potentially from the same wireless modem), then each user is given a different connection identity.

[0130] If the wireless modem cannot be successfully authenticated, then an association response frame with an appropriate reason code will be sent. Different reason codes can be defined to cover each of the possible different reasons for the failure to associate. If it is desired to combine the MAC layer registration with the network layer registration, the association request frame should contain sufficient login information to enable the AP to send a network layer registration packet to the requesting wireless hub. In this case, the AP will not send the association response frame until it receives a further response from the wireless hub.

[0131] If the MAC layer registration is not combined with the network layer registration, then the AP can relay the MAC layer registration to the wireless hub before sending the association response frame. The separation of MAC layer registration and network layer registration is useful if it is desired that the network software be reusable for other physical implementations. Also, if different users are using the same wireless modem to make different connection requests, then the wireless modem may need to make only one MAC layer registration, but may still need to make multiple network layer registrations. If there is only one user for each wireless modem, then combination of MAC layer with network layer registration helps to reduce the number of airlink frames during the registration process.

[0132] Upon receipt of a reconnect message from the remote host PC, a wireless modem reassociates with an access point via the following procedure:

- 50 1. The wireless modem transmits a reassociation request frame to the access point;
2. If the reassociation response frame is received with a status code of "successful", the wireless modem transmits a reconnect success message to the PC;
- 55 3. If the reassociation response frame is received with a status code other than "successful", the wireless modem transmits a reconnect fail message to the PC.

[0133] The access point operates as follows in order to support the reassociation of stations:

into a ratio of  $k_1:1$  between the downlink and uplink traffic.

[0141] The access point buffers are also partitioned into a  $k_2:1$  ratio of downlink to uplink traffic. Again, traffic characterization seems to suggest that a ratio of 4:1 (downlink capacity being 4 times greater than uplink capacity) is appropriate. When the downlink buffer occupancy hits the high threshold, the access point sends an 'Xoff' message to the wireless hub. When the downlink buffer occupancy hits the low threshold (after previously exceeding the high threshold), it sends an 'Xon' message to the wireless hub. When the uplink buffer occupancy hits the high threshold, the access point sets the 'Xon' bit in the frame control field at the time it sends the next broadcast frame to all associated wireless modems. When the uplink buffer occupancy hits the low threshold (after previously exceeding the high threshold), the access point will clear the 'Xoff' bit in the frame control field at the time it sends the next broadcast frame. In addition, a more sophisticated flow control scheme is used by the access point to keep track of the buffer occupancy of each wireless modem (in either direction) and to send an Xon/Xoff MAC frame to a specific wireless modem for a high uplink buffer threshold violation or inform the wireless hub of the appropriate connection ID for a high downlink buffer threshold violation.

[0142] An aspect of the invention is capable of supporting admission control. When a PC user submits a connection request via the wireless modem, the connection request is converted into a network layer registration message that is transmitted across the airlink to the AP. The AP needs to make a decision as to whether to admit this new connection request. The admission control technique can be simple, such as admitting any new connection request if the total number of connections admitted is less than a maximum number. A simple admission control technique cannot guarantee quality of service to all admitted users, however, and may not result in high bandwidth utilization.

[0143] Other admission control techniques may therefore be better than the simple scheme. A specific admission control program may even utilize a combination of several techniques. For example, where each connection request specifies a delay requirement, a bandwidth requirement, and a traffic descriptor, the AP may first compute various performance metrics (e.g. total bandwidth consumed, average delay) in order to determine whether admission of the new connection could cause a failure to meet the Quality of Service of those admitted connections. If the Quality of Service of all admitted connections can be maintained with the admission of the new connection, the new connection will be admitted. Otherwise, the new connection request will be denied. The equivalent bandwidth-based admission technique described by K. M. Rege in "Equivalent Bandwidth and Related Admission Criteria for ATM Systems- A Performance Study," International Journal of Communication Systems, Vol. 2, pp. 181-197 (1994) may be used with minor modifications for handling this problem in a wireless environment. For example, Rege assumes there is only one bandwidth requirement and set of QoS requirements. Here, the method of Rege is extended to support multiple bandwidth requirements and different QoS requirements for uplink/downlink. Adjustment to the bandwidth requirement based on the radio distance (and hence the potential FER that may be experienced) between the wireless modem and the AP is also supported.

[0144] In another example, each connection request specifies the average bit rate required and a traffic burstiness factor. The AP collects information about the number of bytes sent by each connection in either direction for a certain period of time. The AP also measures a burstiness factor for the connection traffic in either direction. Based on this measured information, the AP is able to determine the potential average connection bit rate in either direction (uplink/downlink) and the burstiness factor of each connection. The AP then computes an equivalent number of admitted connections. When a new connection request arrives, the AP calculates whether the new equivalent number of admitted connections exceeds a specified threshold. If the threshold is exceeded, the connection request is denied. Otherwise, it is accepted.

[0145] The measured quantities can be various metrics related to interference. If this is an interference limited system rather than a bandwidth limited system, then, in order to see if the new connection should be admitted, the AP continuously measures a Frame Error Rate (FER) metric for each remote host based on the interference measured. An implementation of this method for admitting new connections based measured quantities in a wireless network is illustrated in Fig. 20. An uplink Frame Error Rate, an average uplink bit rate, a burstiness factor of the uplink traffic, and a packet loss rate are measured 2010 at the base station for each remote host. A downlink Frame Error Rate, an average downlink bit rate, a burstiness factor of the downlink traffic, and a packet loss rate are measured at each admitted remote host 2015, and then the downlink FER is sent 2020 to the base station. This procedure is continuous 2025, allowing all remote hosts currently admitted to send their measured FER to the base station. The reporting process may be either periodic or triggered. In an alternate embodiment, each remote also sends the measured average downlink bit rate, traffic burstiness factor, and packet loss rate to the base station.

[0146] An equivalent bandwidth based on average and peak bit rates of the connection, the burstiness factor of the traffic, and the packet loss rate of each connection is computed 2030 at the base station for each remote host. These computations are continuously updated from new information received from the remote hosts and are used by the base station to compute an equivalent number of connections 2040 already admitted. If a new connection is requested 2045, the base station considers the effect of the average rate and packet loss rate requested by the requested connection and, based on the equivalent bandwidth, computes 2050 whether Quality of Service of all admitted connections

[0155] If only a specific connection is experiencing a high downlink frame error rate, then the access point may elect to disconnect other connections if the connection experiencing bad performance is of a higher priority. For example, when a specific high priority connection is experiencing a high uplink frame error rate, the access point may disconnect other users in order to give more bandwidth to the higher priority connection. If a majority of all associated connections experience high uplink frame error rates, the AP may instead send a congested signal to a wireless hub which can coordinate the actions of other access points, such as by sending signals to these access points to inhibit them from admitting new users and dropping lower priority users.

[0156] There may also be occasions when there is a sudden increase in short bursty messages. Short packets queued up for so long, in either the uplink or downlink queue at the access point, that they exceed the time-to-live value allocated for them will be thrown away, resulting in an increase in packet loss rate due to the processing bottleneck at the access point. Under such an overload situation, the access point may elect to temporarily disconnect some users of a lower priority. Other combinations of the possible actions discussed would also be suitable, the exact combination being decided by the base station depending on the particular congestion conditions observed in the network.

[0157] A particular embodiment of a method for overload control is illustrated in the flowchart of Fig. 21. As seen in Fig. 21, an uplink Frame Error Rate is continuously measured 2110 at the base station for each remote host based on an average uplink bit rate, a burstiness factor of uplink traffic, and a packet loss rate. Similarly, a downlink Frame Error Rate is measured at each remote host 2115 based on the average downlink bit rate, the burstiness factor of the downlink traffic, and the packet loss rate and then each FER is sent 2120 to the base station. This procedure is continuous 2125, allowing all remote hosts currently admitted to send their FER to the base station. If an overload condition exists, flow-control messages are sent between at least one of the remote hosts and the base station in order to control data flow 2130. Packets at the base station having a delay exceeding a time-to-live threshold 2135 are then discarded 2140, and connections with a Frame Error Rates that has exceeded a frame error rate threshold for a specified time 2145 and that have indicated that their connections can be interrupted 2150 are disconnected 2155.

[0158] To obtain a particular quality of service, each connection request contains the following information: bandwidth requirement, delay requirement, a "loss tolerable/non-tolerable" flag, a "service interruption allowed" flag, acceptable packet loss rate, and a traffic descriptor which consists of peak data rate, average data rate, and a potential burstiness factor for each direction, uplink and downlink. For example, a connection that specifies a delay requirement of 20 ms and "loss tolerant" will have its packet thrown away if the message it sends or is supposed to receive sits in the queue at the wireless modem or the access point for more than 20 ms. If the user specifies a delay requirement but classifies itself as "loss non-tolerant", then packets intended for that user will not be thrown away until there is a buffer overflow. The bandwidth requirement, delay requirement, packet loss rate, and the traffic descriptor are all used in the admission control technique.

[0159] A data security feature can be implemented using any of the methods known in the art. One example would be to adapt the Institute of Electrical and Electronics Engineers (IEEE) standard 802.11 wired Local Area Network (LAN) equivalent approach. The wired equivalent privacy (WEP) feature is defined in the 802.11 standard to protect authorized users of a wireless LAN casual eavesdropping. Payload encryption is not be turned on unless the WEP option is turned on. Each service provider assigns a shared key to all users, in addition to a user-unique key. The keys are periodically modified, with the effectiveness of the security feature depending on the length of the key chosen and the frequency with which the key is changed.

40

### Claims

1. A method for overload control in a wireless communications network employing the On-Demand Multiple Access Fair Queuing protocol, said network having a base station and a plurality of remote hosts between which uplink and downlink packets flow, the method comprising the steps, in combination, of:

measuring at said base station an uplink Frame Error Rate, an average uplink bit rate, a burstiness factor of uplink traffic, and a packet loss rate for each of said remote hosts;

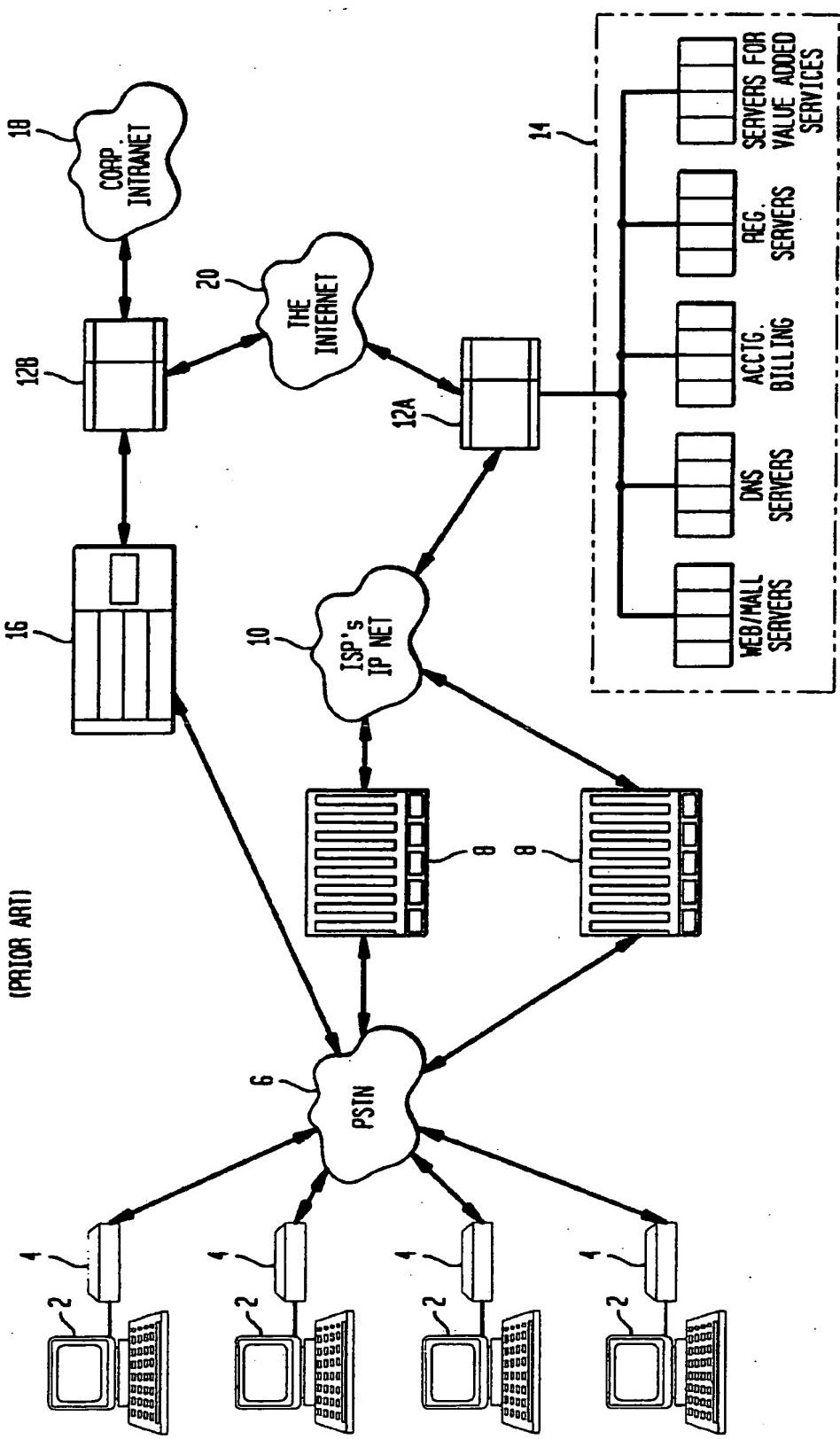
measuring at each of said remote hosts a downlink Frame Error Rate, an average downlink bit rate, a burstiness factor of downlink traffic from said base station, and a packet loss rate;

sending said measured downlink Frame Error Rate from each of said remote hosts to said base station;

detecting at said base station from said Frame Error Rates when a queue overload occurs at least one of said remote hosts or said base station;

- steps of:
- stopping packet flow from at least one of said remote hosts having a Frame Error Rate that exceeds a frame error rate threshold for a specified time and that has previously indicated that disconnection is allowed;
  - stopping packet flow from at least one of said remote hosts having the lowest priority;
  - redirecting packet flow from at least one of said remote hosts having the lowest priority to a nearby base station; and
  - reducing bandwidth share allocated to at least one of said remote hosts that has previously indicated that a variable allocated bandwidth is allowed;
- if said queue overload is in said base station downlink queue, performing at least one of the steps of:
- discarding any of said packets at said base station that have experienced a delay exceeding a time-to-live threshold; and
  - stopping packet flow to said base station from other parts of said network;
- if said queue overload is in said uplink queue at at least one of said remote hosts, performing at least one of the steps of:
- discarding any of said packets at said remote host that have experienced a delay exceeding a time-to-live threshold; and
  - stopping packet flow to said remote host from other parts of said network;
- if said queue overload is in said downlink queue at at least one of said remote hosts, performing at least one of the steps of:
- stopping packet flow from said base station to said remote host; and
  - disconnecting said remote host if it has previously indicated that disconnection is allowed; and
- if a queue overload is detected at a large number of said remote hosts, sending flow control messages to a wireless hub requesting coordination of nearby base stations.
10. The method of claim 9, wherein said step of stopping packet flow from at least one of said remote hosts having Frame Error Rate exceeding said threshold and that has previously indicated that disconnection is allowed is applied to one of said remote hosts having the lowest priority of all said remote hosts having Frame Error Rate exceeding said threshold and that has previously indicated that disconnection is allowed.
11. A method for overload control in a wireless communications network employing the On-Demand Multiple Access Fair Queuing protocol, said network having a base station and a plurality of remote hosts between which uplink and downlink packets flow, said base station and said remote hosts having respective uplink and downlink buffers, the method comprising the steps, in combination, of:
- detecting at said base station when a queue overload condition occurs because of at least one of said remote hosts or said base station, said queue overload condition occurring when the occupancy rate of one or more of said buffers exceeds a high threshold; and
  - performing at least one overload control step, said at least one overload control step being selected from the following:
    - if said remote hosts are of varying priority, disconnecting at least one lower priority one of said remote hosts;

**FIG. 1**  
(PRIOR ART)



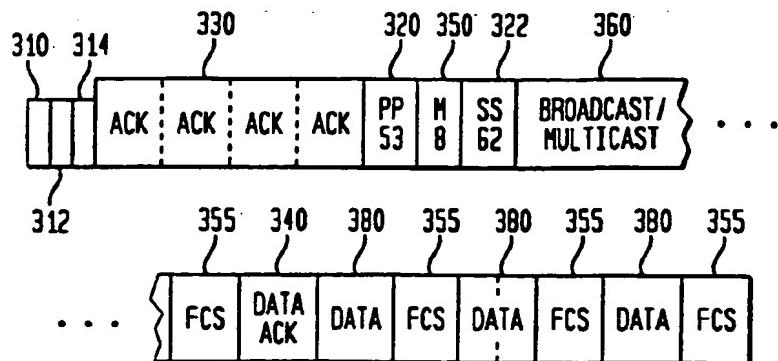
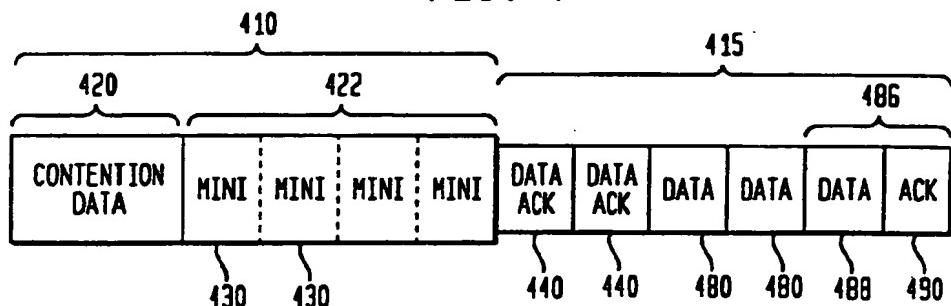
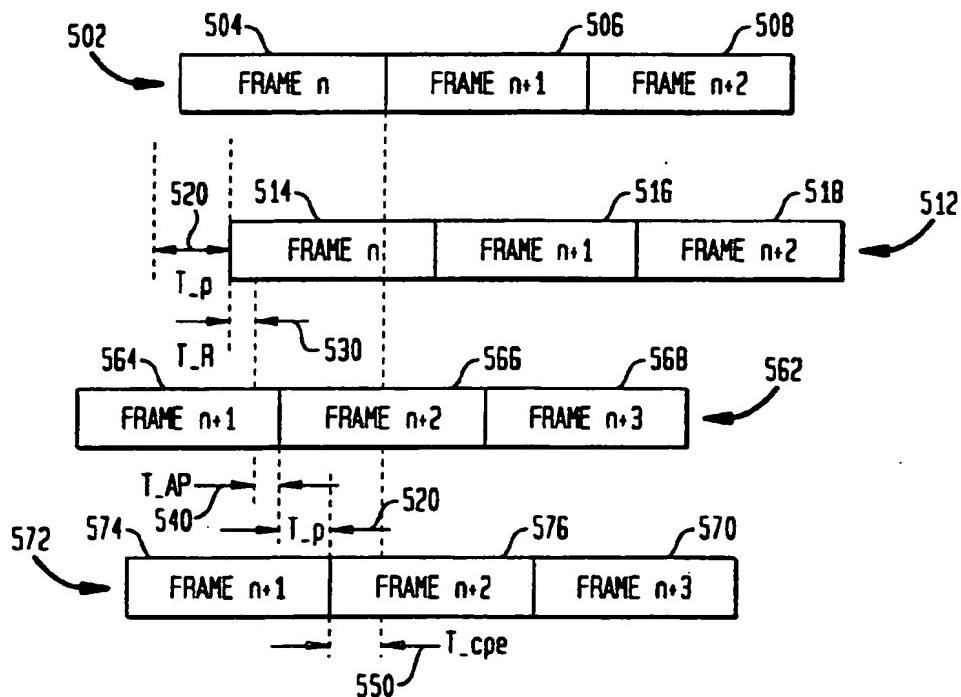
**FIG. 3****FIG. 4****FIG. 5**

FIG. 6C

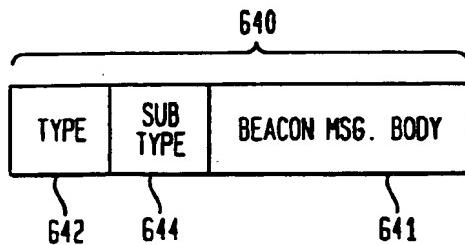


FIG. 6D

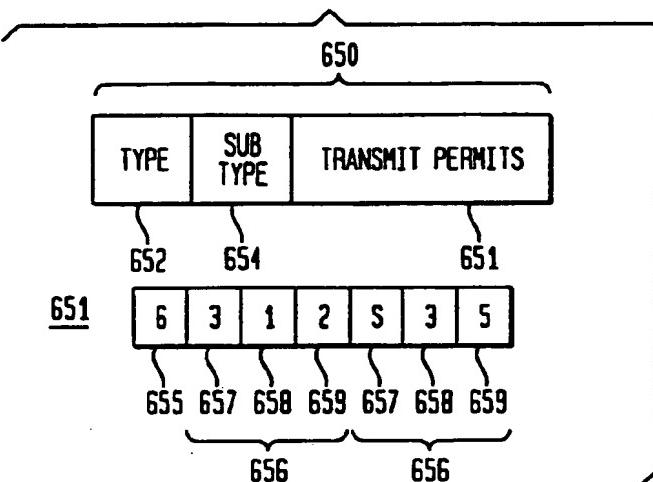


FIG. 6E

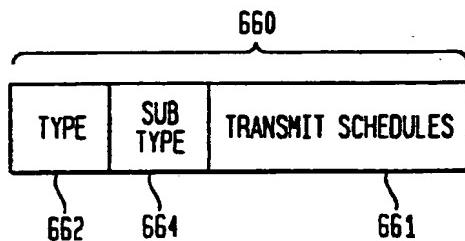


FIG. 6F

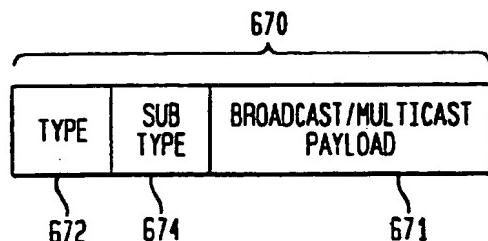
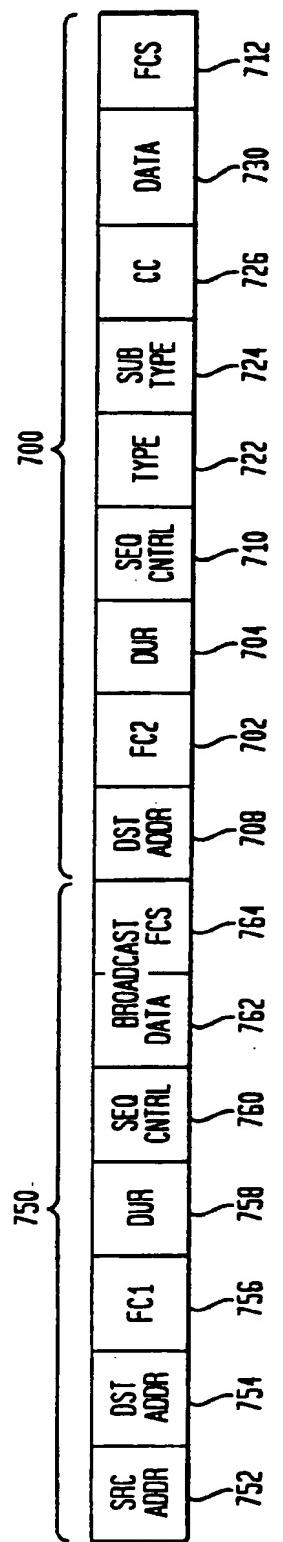
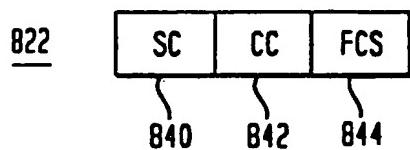


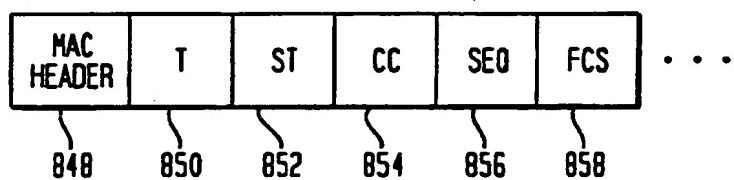
FIG. 7D



**FIG. 8D**



**FIG. 8E**



**FIG. 8F**

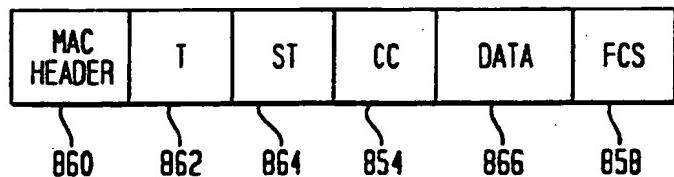


FIG. 9A

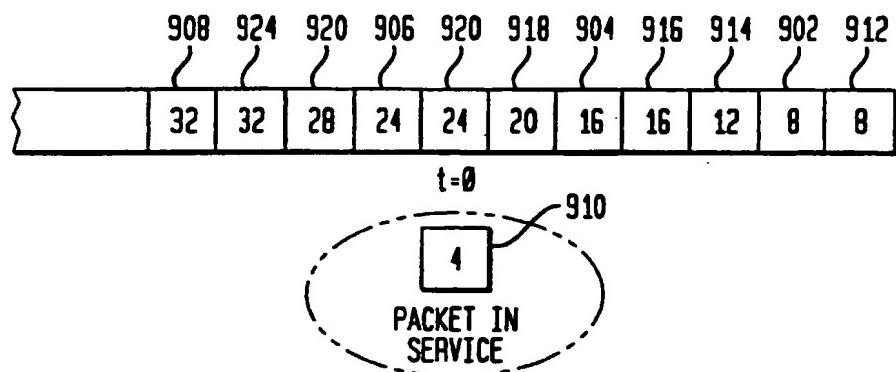
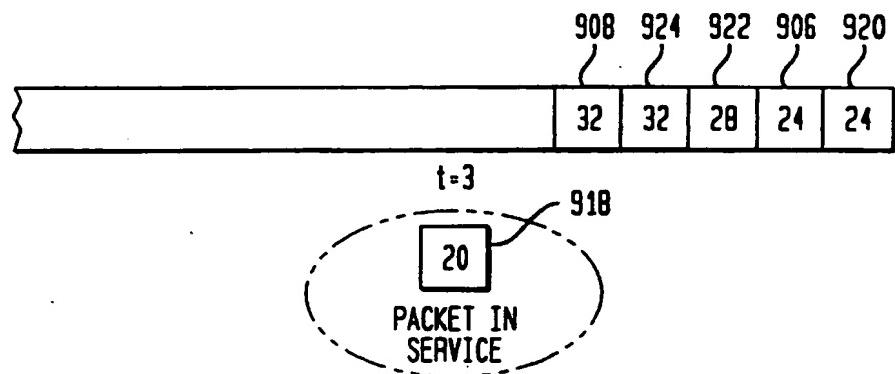


FIG. 9B



**FIG. 10**

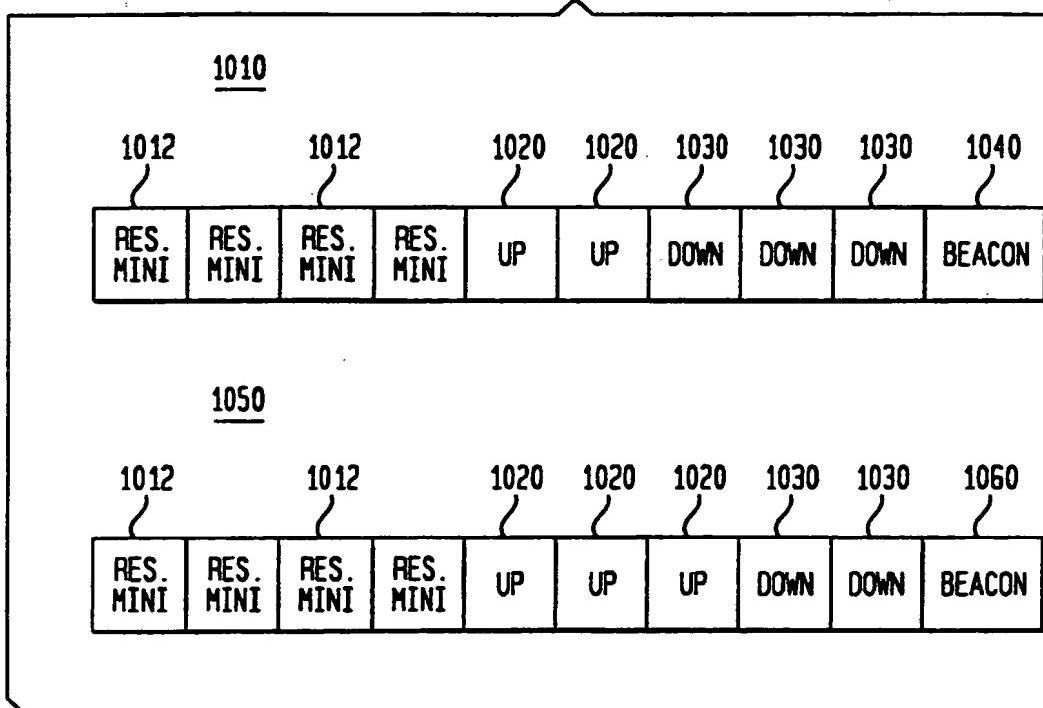


FIG. 12A

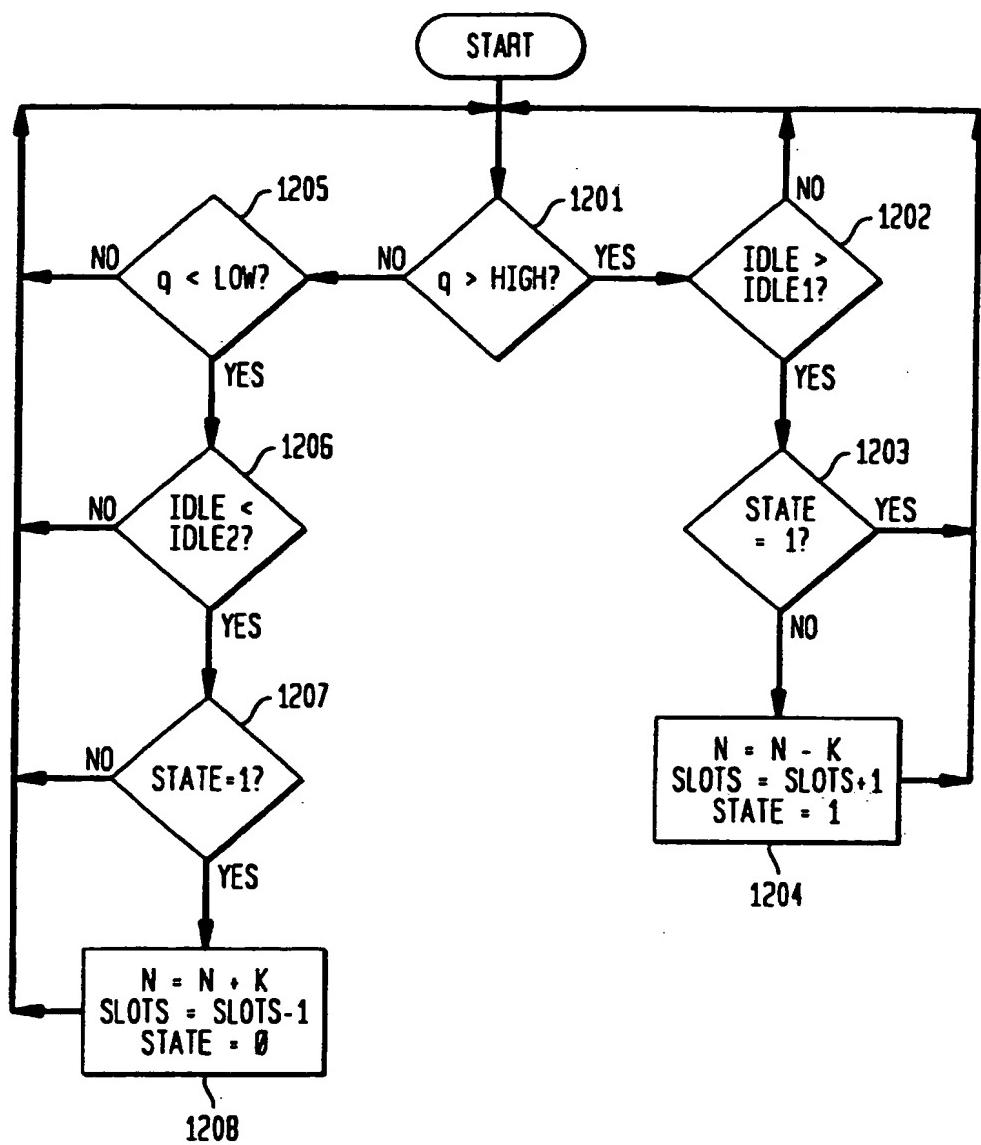


FIG. 12C

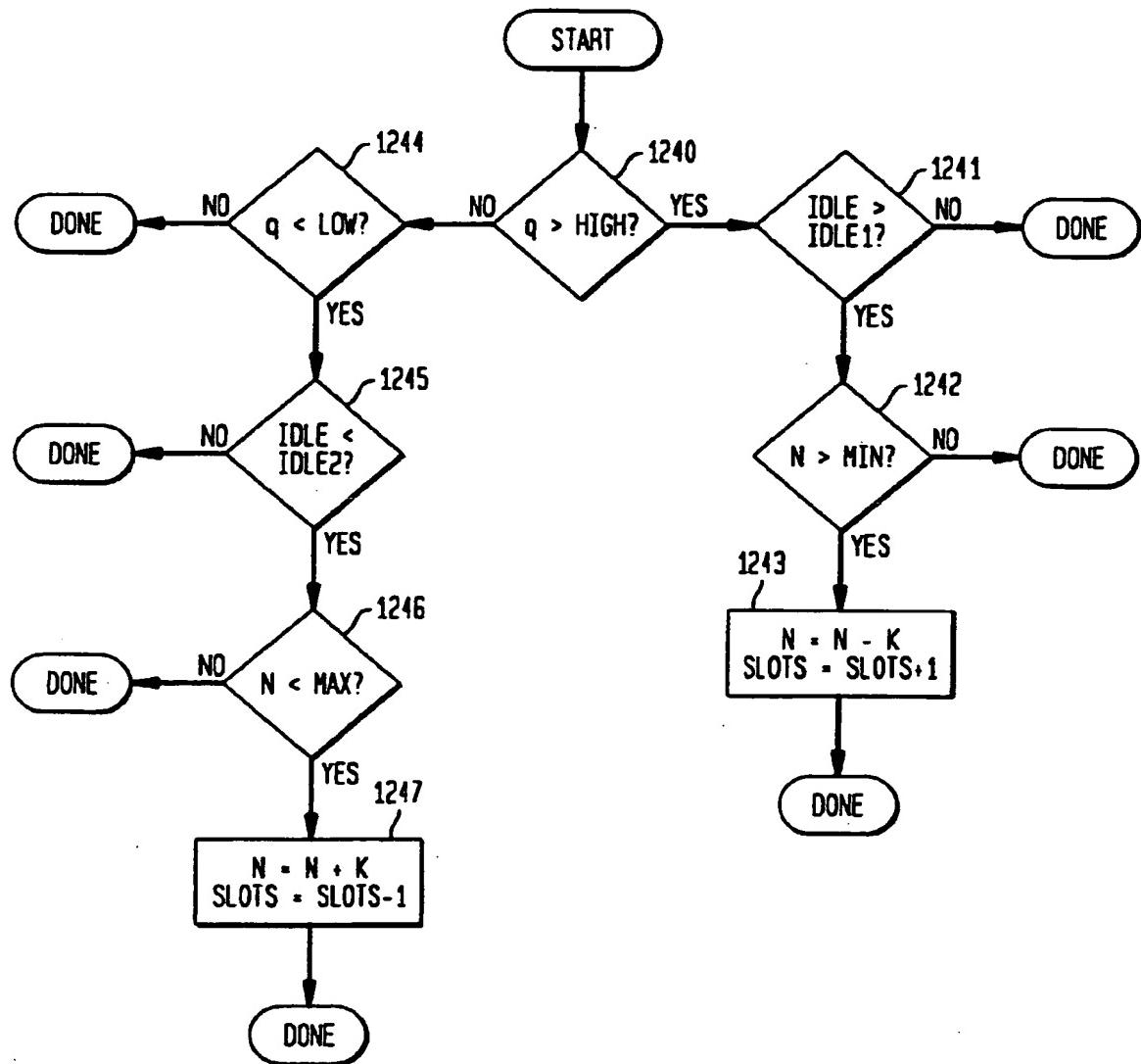


FIG. 13A

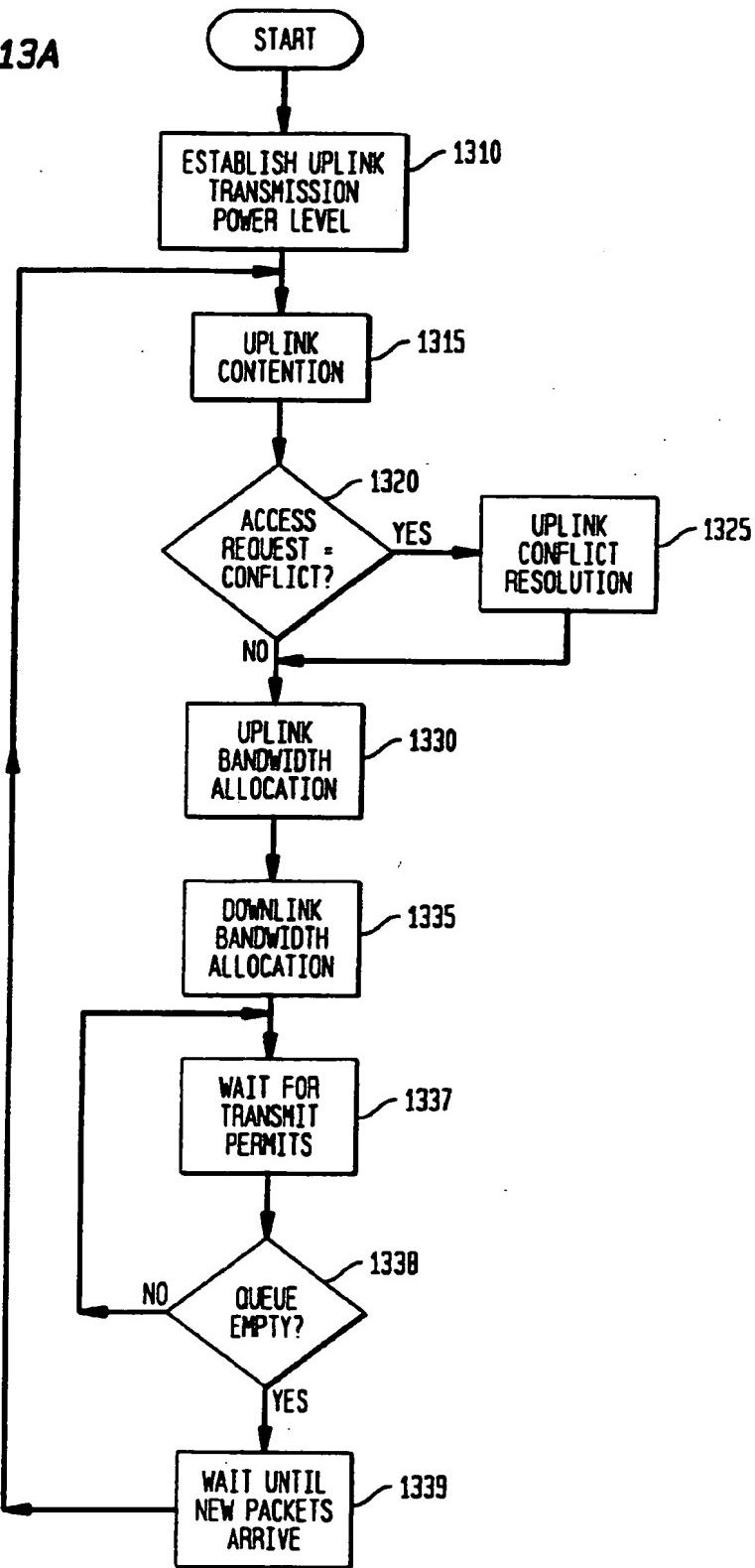


FIG. 14A

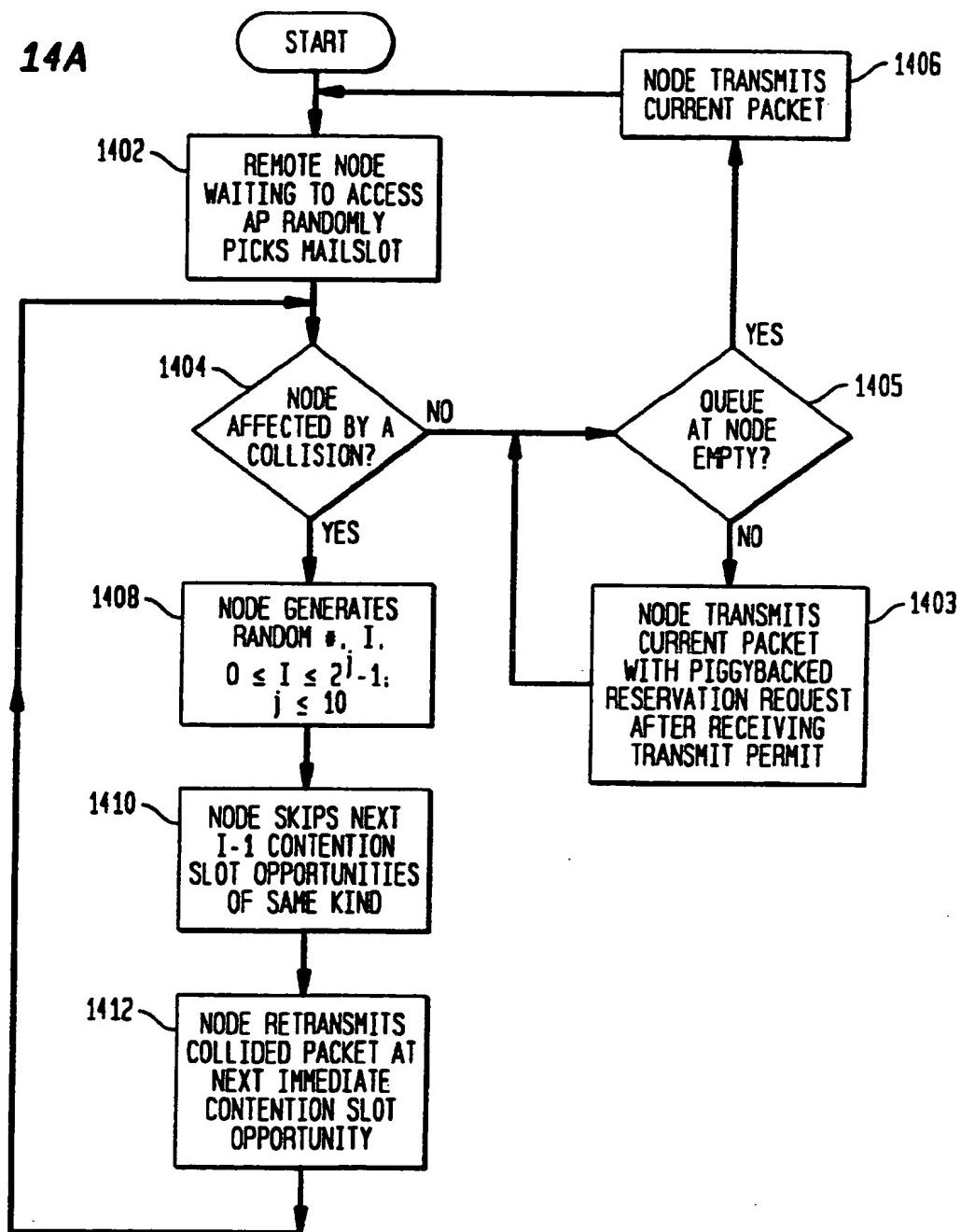


FIG. 14C

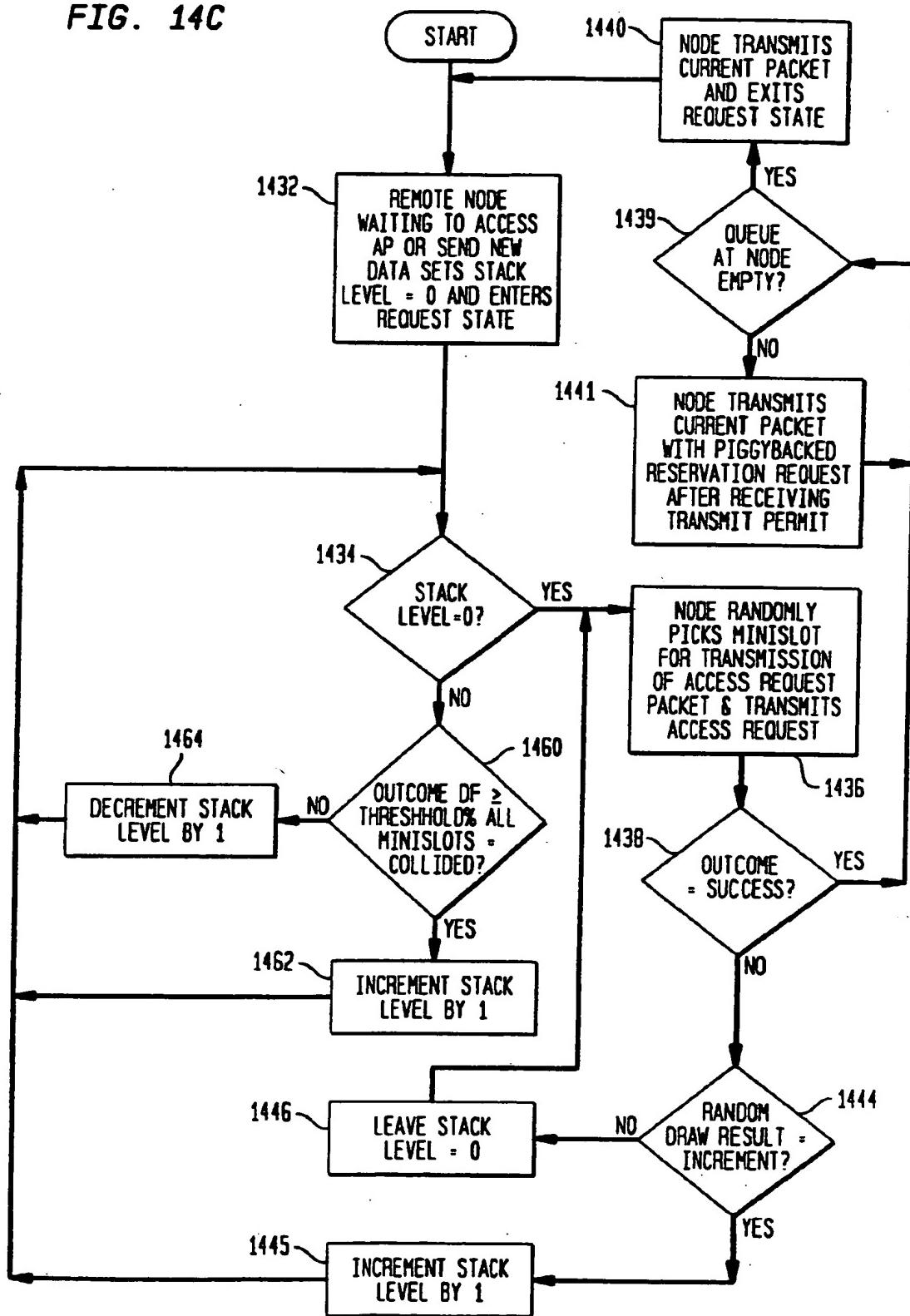


FIG. 15B

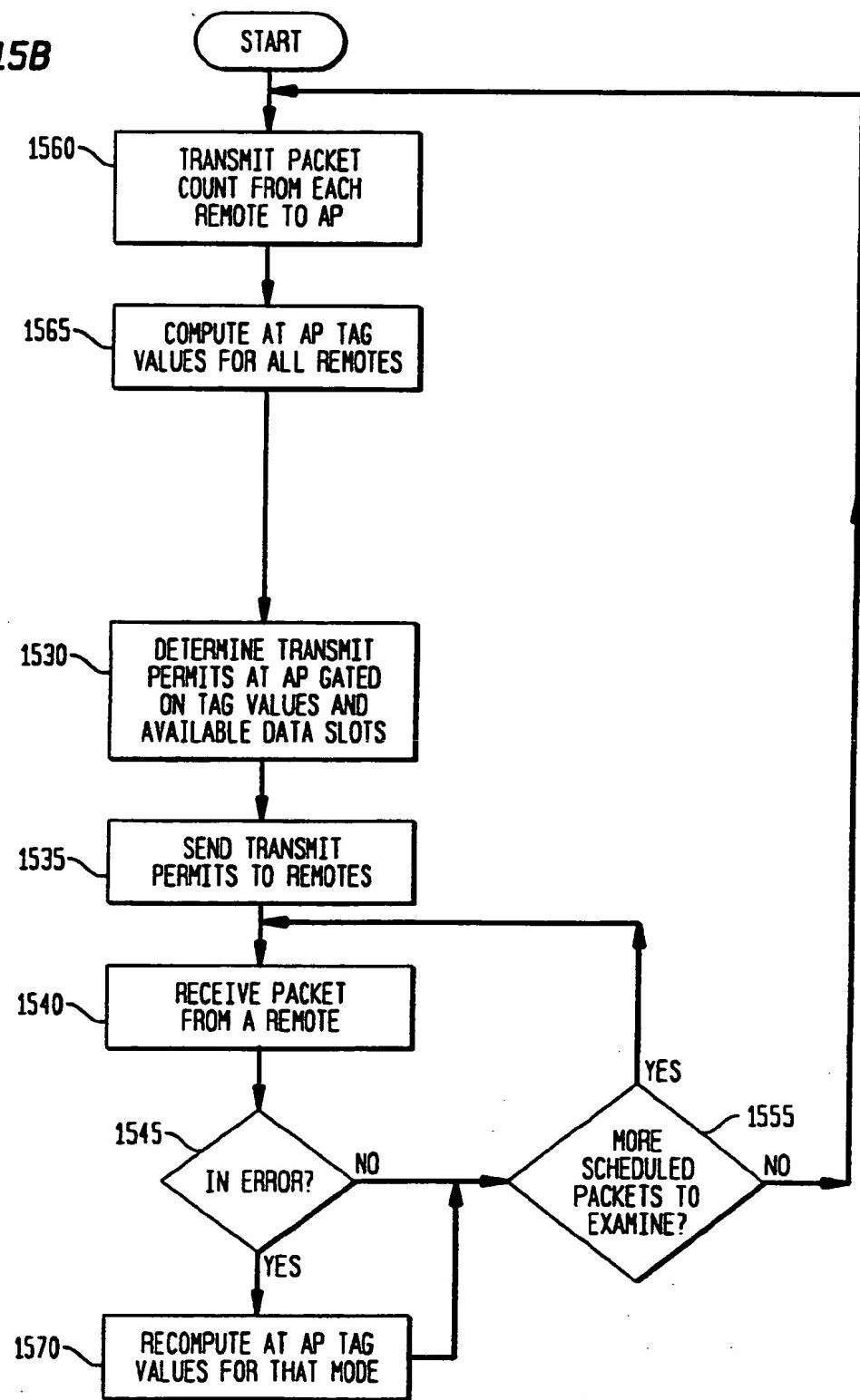


FIG. 17

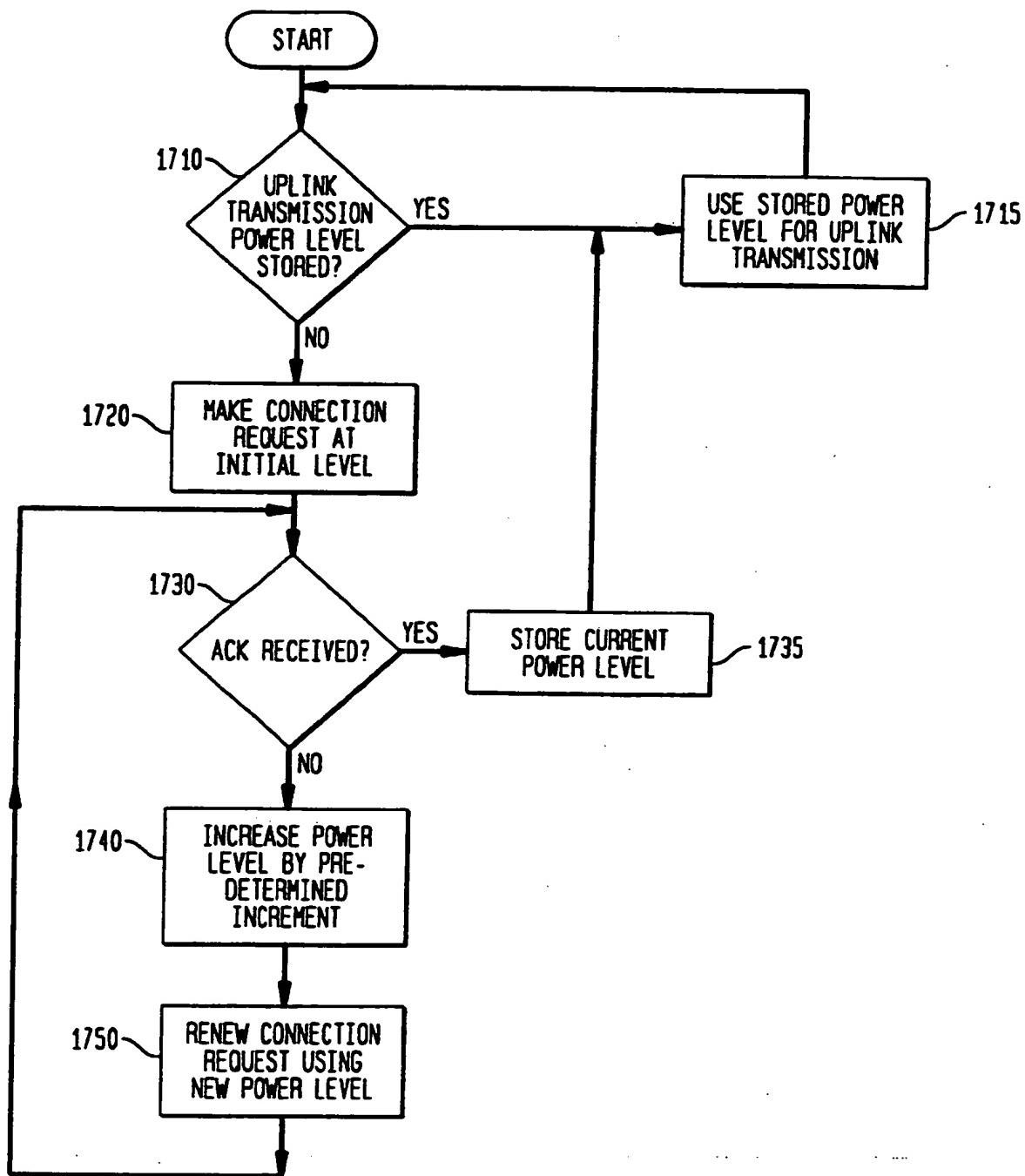


FIG. 18B

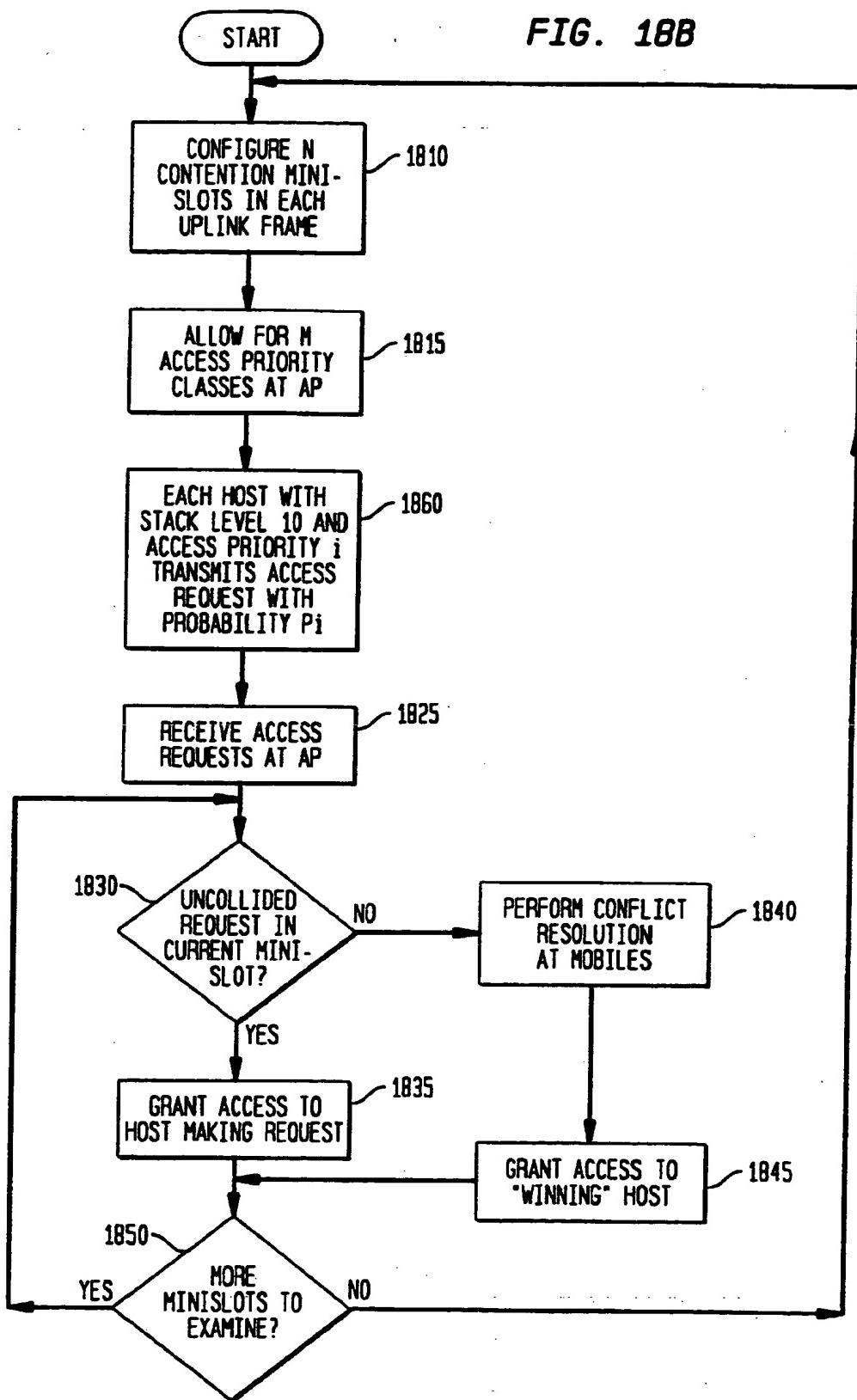


FIG. 20

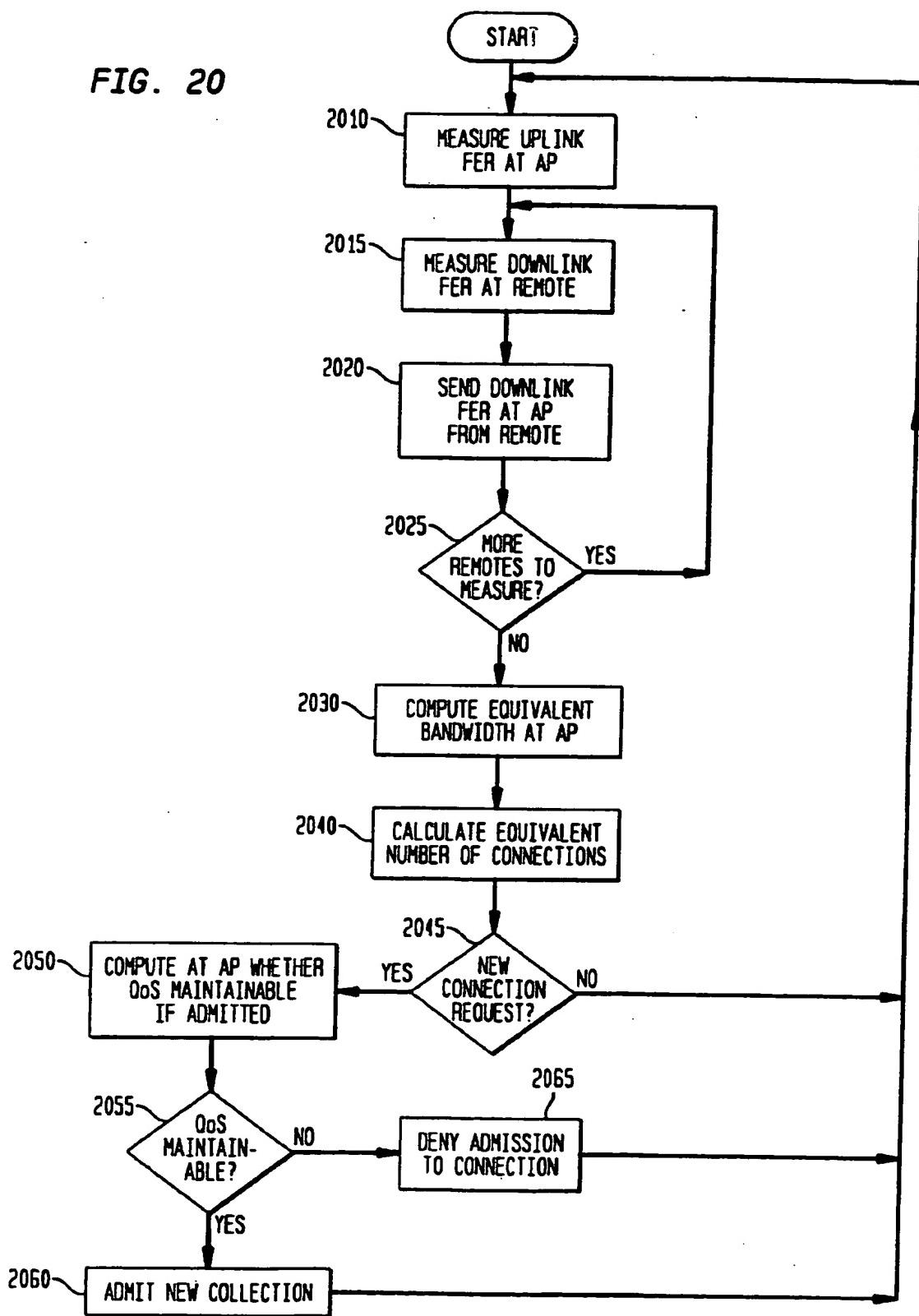
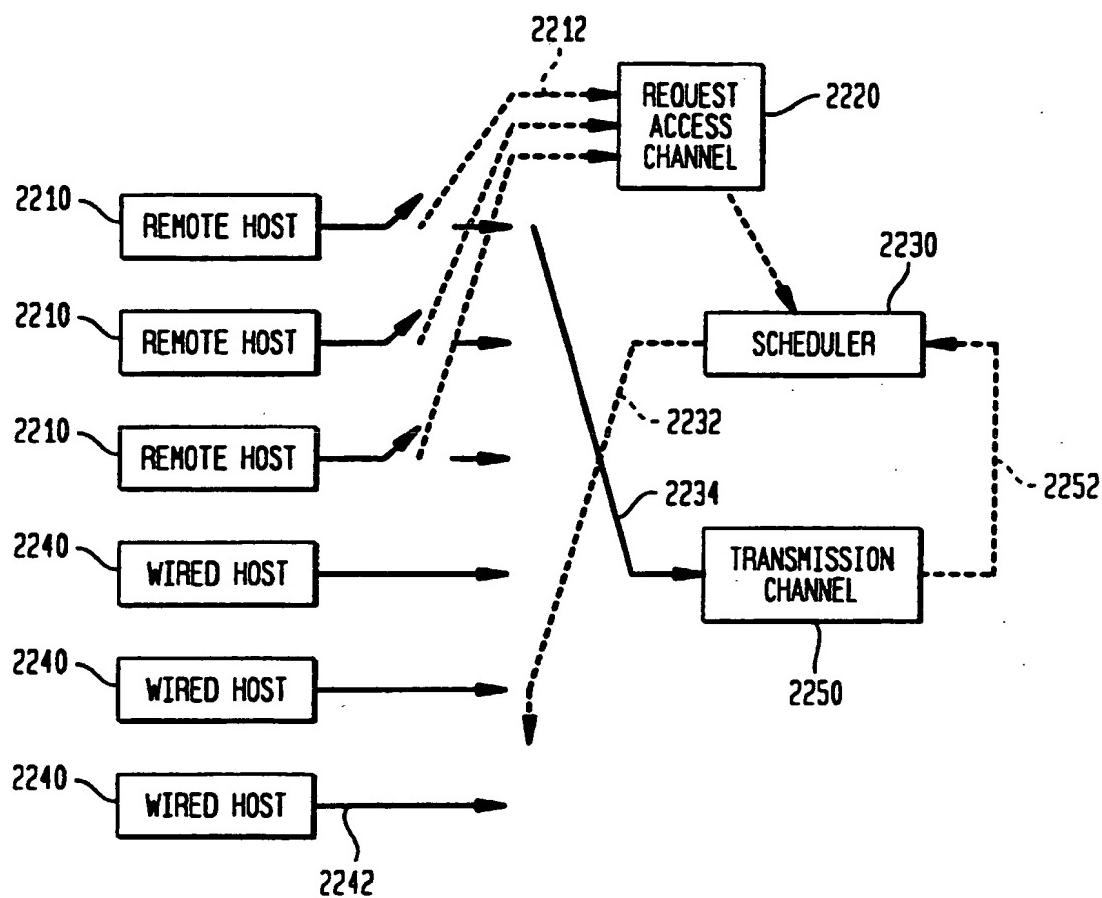


FIG. 22



(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11)

EP 0 912 015 A3

(12)

## EUROPEAN PATENT APPLICATION

(88) Date of publication A3:  
06.05.1999 Bulletin 1999/18

(51) Int Cl. 6: H04L 12/26, H04L 12/56,  
H04L 12/28

(43) Date of publication A2:  
28.04.1999 Bulletin 1999/17

(21) Application number: 98308309.8

(22) Date of filing: 13.10.1998

(84) Designated Contracting States:  
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE

Designated Extension States:  
AL LT LV MK RO SI

(30) Priority: 14.10.1997 US 61790 P  
12.03.1998 US 77741 P  
22.05.1998 US 83759

(71) Applicant: LUCENT TECHNOLOGIES INC.  
Murray Hill, New Jersey 07974-0636 (US)

(72) Inventor: Chuah, Mool Choo  
Eatontown, New Jersey 07724 (US)

(74) Representative:  
Watts, Christopher Malcolm Kelway, Dr. et al  
Lucent Technologies (UK) Ltd,  
5 Mornington Road  
Woodford Green Essex, IG8 0TU (GB)

### (54) Method for overload control in a multiple access system for communications networks

(57) In the method for overload control in a wireless communications network employing On-Demand Multiple Access Fair Queuing, if the downlink/uplink buffer occupancy of the network has exceeded a high threshold, the base station determines if this is caused by a specific remote host or by a group of remote hosts. If caused by a specific remote host, the base station normally sends a flow control signal to the remote host to prevent it from sending more data, but may alternatively elect to disconnect other remotes if the remote experiencing bad performance is of a higher priority. The base station may additionally reduce the bandwidth shares allocated to any remote that have indicated tolerance for a variable allocated bandwidth. If the measured frame error rates for many remote hosts are increasing, then the base station may elect to disconnect those remote hosts that permit service interruption in order that more bandwidth may be allocated to the remaining users. If a majority of all associated remote hosts experi-

ence high uplink frame error rates, the base station may instead send a signal to a wireless hub which can coordinate the actions of other access points. Short packets queued up for so long at the base station that they exceed the time-to-live value allocated will be thrown away. The base station may also or alternatively elect to disconnect some users of a lower priority or redirect them to other nearby base stations that have a lower load. In a particular embodiment, an uplink Frame Error Rate (FER), an average uplink bit rate, a burstiness factor of uplink traffic, and a packet loss rate are measured at the base station for each remote host. Similarly, a downlink Frame Error Rate is measured at each remote host and then each FER is sent to the base station. If an overload condition exists, connections with a Frame Error Rate that has exceeded a threshold for a specified time and that have indicated that their connections can be interrupted are disconnected.

EP 0 912 015 A3



European Patent  
Office

## EUROPEAN SEARCH REPORT

Application Number

EP 98 30 8309

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
P, X	US 5 721 762 A (SOOD PREM) 24 February 1998 * column 1, line 21 - column 6, line 5 * * column 7, line 18 - column 8, line 59 * * column 12, line 21 - column 14, line 32 * * column 15, line 17 - column 16, line 55 * * column 17, line 57 - line 65 * * claims 1-4,6 *	11	H04L12/26 H04L12/56 H04L12/28
A X	& JP 09 215061 A (SHARP) 15 August 1997 * the whole document *	1,9 11	
A	US 5 335 357 A (FENNELL ROBERT D ET AL) 2 August 1994 * column 1, line 9 - line 11 * * column 2, line 26 - line 35 * * column 3, line 34 - line 62 *	1,9,11	
A	WO 95 15644 A (ERICSSON TELEFON AB L M) 8 June 1995 * page 1, line 5 - line 9 * * page 3, line 9 - page 7, line 3 * * page 13, line 15 - page 14, line 17 *	1,9,11	TECHNICAL FIELDS SEARCHED (Int.Cl.6) H04L H04Q
A	WO 95 35002 A (QUALCOMM INC) 21 December 1995 * page 1, line 7 - line 11 * * page 2, line 27 - page 5, line 3 * * page 16, line 9 - page 17, line 22 *	1,9,11 -/-	
The present search report has been drawn up for all claims			
Place of search	Date of completion of the search	Examiner	
THE HAGUE	3 March 1999	Vaskimo, K	
CATEGORY OF CITED DOCUMENTS			
X : particularly relevant if taken alone	T : theory or principle underlying the invention		
Y : particularly relevant if combined with another document of the same category	E : earlier patent document but published on, or after the filing date		
A : technological background	D : document cited in the application		
O : non-written disclosure	L : document cited for other reasons		
P : intermediate document	& : member of the same patent family, corresponding document		



European Patent  
Office

## EUROPEAN SEARCH REPORT

Application Number  
EP 98 30 8309

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
A	EP 0 660 626 A (NOKIA MOBILE PHONES LTD) 28 June 1995 * abstract * * column 1, line 1 - line 12 * * column 4, line 22 - line 57 * -----	1,9,11	
TECHNICAL FIELDS SEARCHED (Int.Cl.6)			
The present search report has been drawn up for all claims			
Place of search	Date of completion of the search	Examiner	
THE HAGUE	3 March 1999	Vaskimo, K	
CATEGORY OF CITED DOCUMENTS		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	
EPO FORM 1500/02 02/82 (PVC01) <small>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</small>			

**ANNEX TO THE EUROPEAN SEARCH REPORT  
ON EUROPEAN PATENT APPLICATION NO.**

EP 98 30 8309

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on. The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

03-03-1999

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP 0599515 A		CA	2105544 A	24-05-1994
		JP	2813119 B	22-10-1998
		JP	6216972 A	05-08-1994
GB 2311912 A	08-10-1997	AU	1664897 A	11-06-1998
		BR	9701573 A	23-06-1998
		CA	2199310 A	04-10-1997
		JP	10042338 A	13-02-1998
US 5539729 A	23-07-1996	NONE		
EP 0660626 A	28-06-1995	GB	2285555 A	12-07-1995
		AU	690932 B	07-05-1998
		AU	8026994 A	29-06-1995
		CN	1116396 A	07-02-1996
		GB	2321576 A, B	29-07-1998
		JP	7298339 A	10-11-1995